Doctoral Thesis

Strategic resource management for power grid operators

Author(s):
Guarisco, Michael Simon

Publication Date:
2011

Permanent Link:
https://doi.org/10.3929/ethz-a-006683206

Rights / License:
In Copyright - Non-Commercial Use Permitted
Strategic Resource Management for Power Grid Operators

A dissertation submitted to
ETH ZURICH

for the degree of
DOCTOR OF SCIENCES

presented by
MICHAEL SIMON GUARISCO
Dipl. Math. ETH
born 31st August 1977
citizen of Thalwil (ZH)

accepted on the recommendation of
Prof. Dr. Hans-Jakob Lüthi, examiner
Prof. Dr. Göran Andersson, co-examiner
Prof. Dr. Markus Zdrallek, co-examiner

2011
Acknowledgments

During my time at the Institute for Operations Research (IFOR) at ETH Zurich, I not only had the possibility to deepen my knowledge in Operations Research but I also had the chance to be part of a joint research project with an industrial partner, RWE Rhein-Ruhr Netzserservice GmbH.

First of all, I would like to express my deepest gratitude to Prof. Hans-Jakob Lüthi, the examiner of my thesis. I greatly appreciate that he gave me the opportunity to write my dissertation on this applied project with RWE and I am very grateful for his continuous, encouraging, and kind support, the numerous fruitful discussions and all his valuable ideas and input that were essential for the progress of my thesis.

I am also very grateful to Prof. Göran Andersson for agreeing to be co-examiner of my thesis. I appreciate his friendly support, the interesting discussions, and his helpful suggestions.

I further owe many thanks to Prof. Markus Zdrallek. He not only agreed to be co-examiner of my thesis, but he also led the project at RWE during two thirds of our collaboration. I am very grateful for all the stimulating and constructive discussions and his valuable suggestions and comments that helped improve this thesis.

I am very much indebted to PD Dr. Marco Laumanns, who supervised my thesis and led the project at IFOR. In numerous meetings, he helped me develop and refine the models. He constantly supported my work and contributed essentially to the success of my thesis. He further proofread this thesis at various stages, for which I am very grateful.

I would like to express my gratitude to the RWE group and in particular to RWE Rhein-Ruhr Netzserservice GmbH for funding and supporting our joint research project, and I would like to thank all the people at RWE who were involved in the project. I greatly appreciate the excellent collaboration and the extremely pleasant and friendly working atmosphere.
I very much enjoyed all our workshops and meetings in Siegen, Neuss, Brühl, Berzdorf, Brauweiler, Bergheim-Paffendorf, and Zurich. In particular, I am very grateful to Gerd-Günther Doege, the former technical director of RWE Rhein-Ruhr Netzserservice GmbH, for initiating and constantly supporting the project. I would like to thank Dr. Stefan Küppers for leading and bringing forward the project during its first part. Again, I would like to thank Prof. Markus Zdrallek for leading the project and for all his valuable input. I would also like to thank Dr. Adolf Schweer for his interest in the project and his support. Very special thanks go to Catharina Friedrich, our main contact at RWE, for the very amicable collaboration, for patiently answering my questions, for providing the input data, and for all the fruitful discussions. I would also like to thank Marcus Stötzel, particularly for providing the input data for the second model and for answering my (electro)technical questions.

I further owe many thanks to Dr. Rico Zenklusen for the numerous stimulating discussions and all his valuable ideas and input, that not only supported the project with RWE but also substantially added to the success of my thesis.

I would also like to thank Christian Balderer for all the fruitful discussions, his valuable input, and his help with the implementation of the first model. I am grateful to Stefan Ravizza who contributed to the project with his Bachelor’s thesis and the implementation of the model for planned work. I would like to thank Maria Moschovitis for analyzing new key performance indicators in her term project. Special thanks go to Samuel Zürcher for his valuable input and the many interesting discussions. In a term project and his Master’s thesis, he laid the basis for the second model.

Very special thanks go to my former and current colleagues at IFOR. I very much appreciate the amicable and sociable atmosphere at IFOR and I enjoyed the many pleasant hours with you, also outside the institute.

Many thanks go to Barbara Gut, Jan Mattmüller, and Charles Mitchell for proofreading parts of the thesis.

I would like to express my deepest gratitude to my parents and my sister for their love and encouragement, to all my friends for their support and friendship, and to my girlfriend Barbara Gut for her great patience and understanding, her continuous support, and her love. Thank you all for believing in me and for always being there for me.
The liberalization of electricity markets has led to an increased cost pressure on power grid operators. As they are also responsible for the reliable supply of electricity, power grid operators aim to optimally balance the opposing objectives of cost and quality of supply. One main aspect of the quality of supply is the continuity of supply, i.e., the availability of electricity to consumers. The continuity of supply strongly depends on the restoration time after incidents in the power grid, which is directly influenced by the availability of (human) resources performing the restoration work. Power grid operators thus try to find an organization of resources that guarantees a high quality of supply at minimum cost. This thesis focuses on strategic resource management for the restoration work after incidents in the power grid. Two models are presented for analyzing the effects of a limited availability of resources on the continuity of supply. The restoration times in both models endogenously depend on the resources currently available.

The first model is a grid operation model that simulates in detail the resources’ activities for restoration after incidents in all voltage levels. An assignment problem is repeatedly solved to decide which incidents are restored by the resources currently available. Results of the model include estimates of the non-availability of supply, which is an indirect measure of the continuity of supply, and of probability distributions of, e.g., the restoration time of incidents. By evaluating various key performance indicators, different organizations of resources can be analyzed and compared. New key performance indicators based on incidents without interruption of supply are suggested. Due to its applicability to real-world instances, the model provides a useful tool to support strategic decisions of power grid operators concerning resource management.

The second model is a component-based model for redundant power grids with exponential failure and repair rates. The repair, however, is only
performed if a resource is available. The optimal assignment of resources to failures is determined by a Markov decision process that minimizes the average expected power not supplied over an infinite time horizon. To cope with the large number of states, an aggregated model is formulated that yields an upper bound on the base model. Even though the effects of a varying number of resources on the continuity of supply can be quantified, the observed effect is only marginal. The continuity of supply is determined much more by the redundancies in the power grid and the failure/repair rates than by the number of available resources. The results of the model confirm the high redundancy in meshed power grids and the very rare occurrence of failures.
Zusammenfassung


Das erste Modell ist ein Netzbetriebsmodell, welches die Aktivitäten der Ressourcen zur Behebung von Störungen in allen Spannungsebenen des Stromnetzes detailliert simuliert. Um zu entscheiden, welche der Störungen von den aktuell verfügbaren Ressourcen behoben werden sollen, wird wiederholt ein Zuordnungsproblem gelöst. Das Modell liefert u.a. Schätzer für die Nichtverfügbarkeit, ein indirektes Mass für die Versorgungszuverlässigkeit, sowie geschätzte Wahrscheinlichkeitsverteilungen, z.B. für die Störungsdauer. Durch die Auswertung von verschiedenen Kennzahlen können verschiedene Ressourcenorganisationen analysiert und verglichen werden. Es werden neue Kenngrössen vorgeschlagen, die auf Störungen ohne Versorgungsunterbrechung basieren. Der simulative Charakter des Modells erlaubt die Anwendung auf reale Probleminstanzen und daher
bietet das Modell ein nützliches Instrument, um strategische Entscheidungen von Stromnetzbetreibern im Bereich Ressourcenmanagement zu unterstützen.

Contents

Acknowledgments iii
Abstract v
Zusammenfassung vii

1 Introduction 1
  1.1 Background and motivation 2
    1.1.1 Liberalization of the European electricity sector 2
    1.1.2 Conflicting objectives for power grid operators 2
  1.2 Goal of the thesis 4
  1.3 Thesis overview and contributions 5

2 Power grid operation and modeling approaches 7
  2.1 Power grid operation 8
    2.1.1 Quality of supply and its regulation 9
    2.1.2 Our focus within grid operation 10
    2.1.3 Power grid and effects of incidents 12
  2.2 Modeling approach and two foci 16
    2.2.1 Two different foci 17
  2.3 Related work 19

3 Detailed grid operation model (focus I) 25
  3.1 Power grid, incidents, and restoration process 27
    3.1.1 Incidents in MV/LV 28
    3.1.2 Incidents in HV/EHV 30
  3.2 Resources and their organization 33
  3.3 Randomness, gain of information, and scenarios 34
    3.3.1 Scenario of incidents 35
<table>
<thead>
<tr>
<th>Content</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4 Assignment of resources to incidents</td>
<td>36</td>
</tr>
<tr>
<td>3.4.1 Preliminary remarks and a special case</td>
<td>36</td>
</tr>
<tr>
<td>3.4.2 Assignment: general case</td>
<td>38</td>
</tr>
<tr>
<td>3.5 Evaluation and key performance indicators</td>
<td>42</td>
</tr>
<tr>
<td>3.5.1 New key performance indicators for incidents without interruption of supply</td>
<td>45</td>
</tr>
<tr>
<td>3.6 Case study</td>
<td>48</td>
</tr>
<tr>
<td>3.6.1 MV/LV grid</td>
<td>48</td>
</tr>
<tr>
<td>3.6.2 Combination of HV/EHV and MV/LV grid</td>
<td>55</td>
</tr>
<tr>
<td>3.6.3 Summary</td>
<td>60</td>
</tr>
<tr>
<td>4 On Markov chains and Markov decision processes</td>
<td>61</td>
</tr>
<tr>
<td>4.1 On homogeneous Markov chains</td>
<td>62</td>
</tr>
<tr>
<td>4.1.1 On discrete-time homogeneous Markov chains</td>
<td>63</td>
</tr>
<tr>
<td>4.1.2 On continuous-time homogeneous Markov chains</td>
<td>69</td>
</tr>
<tr>
<td>4.2 On Markov decision processes</td>
<td>76</td>
</tr>
<tr>
<td>4.2.1 Discrete-time and continuous-time MDPs</td>
<td>77</td>
</tr>
<tr>
<td>4.2.2 Discrete-time MDPs with average expected cost</td>
<td>85</td>
</tr>
<tr>
<td>4.2.3 Solution methods for unichain discrete-time MDPs</td>
<td>91</td>
</tr>
<tr>
<td>4.2.4 Solving unichain continuous-time MDPs with average expected cost</td>
<td>99</td>
</tr>
<tr>
<td>5 Resource-constrained power grid model (focus II)</td>
<td>103</td>
</tr>
<tr>
<td>5.1 Power grid, failures, and repair resources</td>
<td>104</td>
</tr>
<tr>
<td>5.1.1 Power grid model</td>
<td>105</td>
</tr>
<tr>
<td>5.1.2 Failures, failing components and resources</td>
<td>107</td>
</tr>
<tr>
<td>5.2 Base model</td>
<td>111</td>
</tr>
<tr>
<td>5.2.1 State of the system $x$</td>
<td>112</td>
</tr>
<tr>
<td>5.2.2 Control $u$</td>
<td>113</td>
</tr>
<tr>
<td>5.2.3 State transitions</td>
<td>115</td>
</tr>
<tr>
<td>5.2.4 Properties of the Markov decision process</td>
<td>121</td>
</tr>
<tr>
<td>5.2.5 DC load shedding model for the cost</td>
<td>129</td>
</tr>
<tr>
<td>5.3 Aggregated model</td>
<td>135</td>
</tr>
<tr>
<td>5.4 Case study</td>
<td>146</td>
</tr>
<tr>
<td>5.4.1 Power grid, input data, and problem instances</td>
<td>146</td>
</tr>
<tr>
<td>5.4.2 Results of base model for instance $I_1$</td>
<td>151</td>
</tr>
<tr>
<td>5.4.3 Results of aggregated model for instance $I_1$</td>
<td>156</td>
</tr>
<tr>
<td>5.4.4 Results of aggregated model for instance $I_2$</td>
<td>157</td>
</tr>
<tr>
<td>5.4.5 Effect of traveling</td>
<td>161</td>
</tr>
</tbody>
</table>
5.4.6 More sensitive results for an (unrealistic) instance $I_1^{10^4000}$ ........................................... 162
5.4.7 Summary .................................................. 163

6 Conclusions .................................................. 165
6.1 Conclusions .................................................. 166
6.2 Outlook ...................................................... 168

A Additional material ........................................ 171
A.1 More on discrete-time homogeneous Markov chains .......... 172
A.2 Further remarks on the DC load shedding model ............. 177
A.3 Calculation of failure and repair rates ........................ 186

B Proofs .......................................................... 189
B.1 Proofs of Chapter 3 .......................................... 190
B.2 Proofs of Chapter 5 .......................................... 192

Bibliography .................................................... 235
Chapter 1

Introduction
1.1 Background and motivation

1.1.1 Liberalization of the European electricity sector

The 1996 European Union directive concerning common rules for the internal market in electricity [Dir97] formed the basis for liberalizing the electricity sector in Europe. By breaking up the monopolistic positions and vertically integrated structures of electricity suppliers, the electricity sector was gradually opened for competition. The aim was to increase the efficiency in the generation, transmission, distribution, and sales of electricity while assuring a secure and reliable supply. However, the power grid, being considered a natural monopoly (because building competing grids would be noneconomical),\(^1\) is excluded from competition.

The translation of this directive into national law in each of the member countries of the European Union varied considerably both in manner and pace.\(^2\) Based on the experiences gained, the rules and measures of the directive were revised in 2003 and 2009.\(^3\) The directive prescribes separating the activities of transmission and distribution of electricity from those of generation, trading, and sales (so-called *unbundling*). The operators of the power grid not only have to assure a secure and reliable operation of the power grid but are also obliged to ensure a non-discriminatory access to the power grid for all market participants. Consumers profit from the liberalization by the right to choose their supplier of electricity.\(^4\)

In each member country, a regulatory authority (which we refer to as *regulator*) had to be designated that ensures the efficient functioning of the market. In particular, the regulator approves the grid charges (i.e., the charges for the usage of the power grid and the provision of system services) and may provide standards for the quality of electricity supply.

1.1.2 Conflicting objectives for power grid operators

Power grid operators are facing two conflicting objectives as they are responsible for a high quality of supply, but at the same time the regulation

---

\(^{1}\) cf. [SBFN08, p. 2, 41].
\(^{2}\) cf. [SBFN08, p. 19].
\(^{3}\) cf. [Dir03] and [Dir09], respectively.
\(^{4}\) Non-household customers (including producers and traders) since July 2004 and all customers since July 2007, cf. [Dir03].
of grid charges gives strong incentives for cost reduction. For example, the regulator may fix caps on the revenues of a power grid operator within a regulatory period. By exceeding the cost reduction targeted by the regulator, grid operators can make an additional profit in the current regulatory period. In the subsequent period, the revenue caps are based on the costs of the most efficient grid operator, which again increases the incentives for further cost reductions.\footnote{Such an incentive-based regulation of grid charges is used in Germany since the beginning of 2009, cf. [SBFN08, p. 52, 62].}

To avoid the incentives for cost reduction negatively affecting the quality of supply, regulators may impose quality standards. These standards can be accompanied by a penalty/incentive scheme in case the standard requirements are not met or exceeded. Consequently, grid operators try to find an optimal balance between costs and quality of supply.

The tasks of grid operation (e.g., network control, maintenance of the power grid, restoration work after incidents in the power grid, etc.) require human resources and technical equipment. Resource management is thus a major challenge for grid operators, in particular to find an organization (spatial and temporal availability) of resources that not only minimizes total costs but also guarantees the required quality standards.

One main aspect of the quality of supply is the continuity of supply, i.e., the availability of electricity to consumers, which is usually measured by the duration, extent, and frequency of supply interruptions. Interruptions of supply may be planned by the grid operator, e.g., for maintenance work, or can occur unexpectedly due to (stochastic) incidents in the power grid, e.g., due to technical failures or weather effects. In medium- and low-voltage grids, which are generally operated as radial networks, most

---

**Figure 1.1** Grid operators’ conflict: quality of supply vs. cost.

---
incidents result in an interruption of supply and need supply restoration, which in turn requires resources on site. A temporary non-availability of resources delays the restoration and thus directly influences the quality of supply. By contrast, in the high- and extra-high-voltage power grid, most incidents do not entail an interruption of supply due to the redundancies, its meshed operation and the degree of automation and remote control. Nevertheless, the failed equipment generally needs to be repaired and the availability of resources has a strong impact on the reliability of the system. As the redundancy in the power grid is reduced after an incident, a delay in the repair affects the duration of the grid being in a state of risk. Additional incidents might eventually lead to a (large) interruption of supply.

The strong impact of the availability of resources on the quality of supply confirms the importance of resource management for power grid operators. For strategic decisions on resource management, power grid operators need to analyze and quantify the effect of a limited availability of resources (and of delays in restoration) on the quality of supply. A schematic view of the relations between a grid operator, the regulator, the quality of supply, and the cost is given in Figure 1.1.

1.2 Goal of the thesis

The goal of this dissertation is to develop, implement, and apply models and methods that support strategic decisions of power grid operators concerning resource management, i.e., the organization and deployment of (human) resources. These decisions try to optimally balance the opposing objectives of cost and quality of supply. We focus on resource management for the restoration work after incidents in the power grid as this type of work has a large impact on the quality of supply due to its stochastic nature.

Compared to most of the existing work, more emphasis is put on the detailed modeling of a limited number of resources to better capture the dependency of restoration times on the availability of resources. The developed models and methods should

- reproduce the relevant processes of grid operation on an adequate level of detail,
- model the effects of the (non-)availability of resources on the quality
of supply endogenously, and

- allow the analysis and quantification of the effects of a limited availability of resources and of different organizations of resources on the quality of supply.

1.3 Thesis overview and contributions

In Chapter 2 we motivate our focus within grid operation, namely the restoration work after incidents performed by a limited number of (human) resources. We describe the different structural properties of power grids and the effects of incidents on the continuity of supply, in particular with respect to a (temporary) non-availability of resources. Then we introduce our modeling approach and show the different characteristics of the two models that are presented in this work. Finally, we review and discuss the related literature.

In Chapter 3 we present the first model, a grid operation model that simulates the organization and availability of (human) resources and the restoration processes after incidents in the power grid in detail. Incidents with and without interruption of supply are considered. For incidents without an interruption of supply, a new measure, the so-called power-at-risk, is introduced that quantifies the risk of potential interruptions of supply in the event of additional incidents. To decide which incidents are worked on in the event of a shortage of resources, an assignment problem is solved which uses the marginal restoration efficiency (“power to be restored per time unit”) as decision criterion. By incorporating many operational details, the model allows the analysis and quantification of the effects of limited resources on the quality of supply. New key performance indicators, related to incidents without interruption of supply, are suggested. Due to the applicability to real-world problem instances, the results of the model can be used – and in fact have been successfully used by our industrial partner, RWE Rhein-Ruhr Netzservice GmbH (Siegen, Germany) – to support strategic decisions concerning the organization of resources of a power grid operator. In a case study based on real-world data, the effects of different organizations of resources on the continuity of supply are illustrated.

In Chapter 4 we summarize the theory of Markov chains and Markov decision processes that form the basis of the second model. The summary is restricted to finite sets of states and controls as this is sufficient for the
second model. We particularly highlight the reformulation of continuous-time models as discrete-time models.

In Chapter 5, we introduce the second model, a power grid model, with particular focus on high- and extra-high-voltage grids, with components whose restoration times after incidents depend on the current availability of a limited number of (human) resources. In contrast to the first model, the emphasis lies more on the power grid and the dynamic effects of incidents than on the operational processes. The assignment of resources to failed components in the event of a shortage of resources is optimized with respect to the long-term average power not supplied in a Markov decision process. The structural properties of this process are analyzed. To cope with the large number of states, an approximate model based on the aggregation of states is formulated that provides an upper bound on the base model. Both the base model and the aggregated model are applied in a case study based on a real high-voltage power grid. The results confirm the high reliability of meshed power grids with redundancies. Again, the model allows to quantify the effect of limited resources on the quality of supply. However, the observed effect is only marginal, which is mainly due to the fact that incidents are very rare and the redundancy in these grids is high.

In Chapter 6, we conclude the thesis and give an outlook on possible further investigations.

In Appendix A, we provide some additional material on discrete-time homogeneous Markov chains (Appendix A.1), on the DC load shedding model (Appendix A.2), and on the calculation of the failure and repair rates of the second model (Appendix A.3). In Appendix B, we state the proofs that are mentioned in the main text.
Chapter 2

Power grid operation and modeling approaches
2.1 Power grid operation

We consider a power grid operator who is responsible for the secure and reliable operation of a power grid. The power grid consists of a network of generators, power supply lines, transformers, loads/consumers, etc. and is part of a power supply system where electrical power is generated and distributed to supply consumers.

Besides (remote) network control, grid operation includes the planning, building, inspection/maintenance, and renewal of the power grid as well as restoration work after incidents in the power grid. In this work, an incident or a failure refers to an unexpected deviation of one or several electrical components from the normal operating conditions. Incidents can be caused by weather, overload, technical failures, construction work, etc. Incidents may or may not lead to an interruption of supply depending on the construction and operation of the affected power grid. The different types of power grids and the effects of incidents are discussed in Section 2.1.3. Interruptions of supply may also be planned by the grid operator, e.g., if parts of the power grid have to be disconnected to perform maintenance work.

To perform his tasks, the grid operator needs own and external human resources as well as technical equipment. He has to establish an organizational structure, in particular a number of resources with different qualifications, and decide on an operational policy, i.e., on how to assign the available resources to the various tasks. Due to the random nature of incidents in the power grid, the demand for resources is stochastic and the operational policy has to cope with this uncertainty.

An important aim of the grid operator is to find the “best” organizational structure and operational policy, e.g., an organizational structure and an assignment policy that not only minimize total costs but also assure a high quality of supply and are guaranteed to satisfy all regulatory requirements.

In Section 2.1.1 we review the different aspects of quality of electricity supply and its regulation. Then in Section 2.1.2 we specify our focus within grid operation in this work. Finally, in Section 2.1.3, we describe the characteristics of power grids and the effects of incidents, particularly in relation to a limited number of resources.

1 Adapted from [VDN07, p. 19].
2.1.1 Quality of supply and its regulation

The quality of (electricity) supply includes three main aspects: continuity of supply, voltage quality, and commercial quality. The continuity of supply refers to the availability of electricity. Continuity of supply is also referred to as reliability of supply. The evaluation of the continuity of supply is based on the characteristic data of the interruptions of supply (e.g., duration and affected power) within a given time period, typically one year. The continuity of supply at system level is assessed by so-called continuity or reliability indices, e.g., by the system average interruption duration index (SAIDI), which is the average amount of time during which each consumer is not supplied within one year. For the calculation of the reliability indices, several distinctions are made, cf. [CEE08, Section 2.1]. Usually long (> 3 minutes) and short (≤ 3 minutes) interruptions are distinguished. In addition, the indices are calculated separately for planned and unplanned interruptions. The acceptance of planned interruptions by the consumers is higher because consumers can adapt to planned interruptions as long as they are adequately announced. Usually, (unplanned) interruptions due to exceptional events (e.g., exceptional weather conditions) are further distinguished.

Regulators often specify target levels for one or several reliability indices and set up an incentive/penalty scheme for achieving a better/worse continuity of supply, e.g., by linking the allowed revenues of the grid operators to the realized continuity of supply. As the reliability indices only control the continuity of supply on average, regulators additionally impose guaranteed standards for each consumer, e.g., that the supply after an unplanned interruption has to be restored within 18 hours. If the required standard is not

---

2 cf. [CEE01, p. 3].
3 cf. [CEE08, p. 1].
4 cf. [EUR03, p. 28].
5 cf. [CEE08, p. 6].
6 cf. [IEE04]. For the precise definition of SAIDI, we refer to Section 3.5.
7 Sometimes short interruptions of up to a few seconds (so-called transient interruptions) are further distinguished, cf. [CEE08, p. 8].
8 cf. [CEE01, p. 24].
9 cf. [CEE05, Sections 2.1, 2.3].
10 cf. [CEE05, Sections 2.1, 2.4].
attained, the affected consumer receives a compensation payment. A summary of regulatory standards for the continuity of supply in European countries is provided in [CEE05, Chapter 2]. In the long term, the continuity of supply is also related to the security of supply that ensures the (long-term) balance between user demand and availability of electricity.\footnote{cf. [EUR03, p. 4].}

Voltage quality covers the technical aspects of electricity, e.g., the stability of the magnitude and frequency of voltage over time.\footnote{cf. [EUR03, p. 4].} Voltage quality is assured by setting standard values and limits for the characteristics of the voltage. In Europe, these standards are based on the European norm [CEN10].

Commercial quality concerns the quality of the customer service, e.g., the handling time of customers’ requests.\footnote{cf. [CEE08, p. 107].} The regulation of commercial quality mainly follows two approaches: guaranteed standards and overall standards.\footnote{cf. [CEE08, p. 109].} Guaranteed standards assure certain services for each individual customer, e.g., a customer’s complaint has to be answered within 15 working days.\footnote{cf. [CEE08, p. 116].} If the required quality level is not met the customer usually has to be compensated. On the other hand, overall standards control the quality of a service for a whole group of customers and are usually specified by a quantile, e.g., 90\% of new customers must be connected to the network within 30 days.\footnote{cf. [CEE08, p. 109].} An overview of standards for commercial quality used across Europe can be found in [CEE08, Chapter 4 and Annex 3].

\subsection{2.1.2 Our focus within grid operation}

The focus of this work is on strategic resource management for power grid operators and in particular on the effects of a limited availability of resources on the continuity of supply. We restrict our attention to
restoration work after incidents.\textsuperscript{17} Even though the restoration work after incidents only amounts for about 10\% of the yearly volume of work,\textsuperscript{18} it is this type of work that mainly affects the continuity of supply due to its stochastic nature. Having occurred unexpectedly, the incidents force the grid operator to react at short notice and, if the incident has caused an interruption of supply, the interrupted power has to be restored as quickly as possible. It is thus especially after incidents that the adequacy of an organization of resources is revealed, in particular whether enough of the right resources are available.

The ability of an organization to act quickly is reflected in the restoration times and thus, for incidents with an interruption of supply, directly influences the continuity of supply. The \textit{(total) restoration time} is the duration from the occurrence of an incident until its repair. It includes

- the time to register the incident and create a work order,
- possible delays if the work order cannot be assigned due to the unavailability of resources for the restoration,
- the travel time of the resources to the location of the incident, and
- the time for the actual restoration work on site (e.g., the repair of the component).\textsuperscript{19}

Throughout this work, we assume there is no bottleneck in the control center for registering the incidents or coordinating the restoration processes. In particular, we assume that a work order is created as soon as an incident has occurred. However, the restoration times for the incidents strongly depend on the availability of the resources that perform the restoration work. The effects of a non-availability of resources after incidents is discussed in more detail in Section 2.1.3 in relation to the different structural forms of power grids.

In this thesis, we further focus on the \textit{human} resources of the grid operator that perform the restoration work. In the following, a \textit{resource} thus always refers to a human resource, e.g., a field service employee, etc.

\textsuperscript{17} In [FGLZ09], we describe a more “complete” model for the organization of grid operation that combines the grid operation model for unplanned work (i.e., restoration after incidents) that is described in Chapter 3 with a model for the optimization of \textit{planned} work such as maintenance, building, and renewal of the electrical equipment. The latter model was also analyzed in [Rav07].

\textsuperscript{18} cf. [Küp05, p. 3].

\textsuperscript{19} cf. [ZDT98].
2.1.3 Power grid and effects of incidents

The power grid is divided into the transmission system and the distribution system. The transmission system is designed for the international and inter-regional transport of electrical energy and has a nominal voltage of 380 kV or 220 kV (extra-high-voltage). We also refer to the transmission system as extra-high-voltage (power) grid. The aim of the distribution system is to distribute the electrical energy to the consumers. The distribution system encompasses three subsystems with different nominal voltages: the high-voltage (power) grid (110 kV), the medium-voltage (power) grid (10 kV or 20 kV), and the low-voltage (power) grid (230 V or 400 V). The link between the different voltage levels is established by transformer substations, where the power is transformed from extra-high- to high-voltage, from high- to medium-voltage, or from medium- to low-voltage. Throughout this thesis, we use the abbreviations EHV for extra-high-voltage, HV for high-voltage, MV for medium-voltage, and LV for low-voltage.

Depending on the voltage level and the load density, different structures of power grids can be found. In the sequel, we summarize the most important structural forms and characteristics of power grids. For a more detailed description, we refer to the literature, e.g., [HDS07, Section 3.2] or [SR08, Chapter 5]. The simplest structure is the (branched) radial network, where each load is supplied by exactly one “path” from the feeding transformer substation (cf. Figure 2.1(a)). The main disadvantage of radial networks is that in the event of an incident, the consumers on the same branch suffer from a supply interruption. By adding additional connections between different branches of a radial network, a ring network (cf. Figure 2.1(b)) or a branched ring network (cf. Figure 2.1(c)) is obtained. Generally, switching devices (which we refer to as switches) are installed along the rings that allow, e.g., to isolate a failed piece of equipment and thus to limit the number of consumers affected by an incident (cf. Example 2.1.1 below). The most complicated structure is the meshed grid consisting of several feed-ins and multiple connections (cf. Figure 2.1(d)), i.e., there exist several possibilities to supply a given point in the grid. Due to these redundancies, one (or even several) incidents can occur without an interruption of supply.

---

20 The indicated nominal voltages in this section are the usual values in Germany, cf. [HDS07, Section 3.2].
21 cf. [HDS07, p. 78].
22 Our summary is based on [HDS07, Section 3.2] and [SR08, Chapter 5].
2.1 Power grid operation

Figure 2.1 Different structures of power grids.

We refer to manual or automatic changes of the configuration of the power grid, e.g., opening or closing electric circuits by the use of switches, as *switching operations*. The following example illustrates the use of switching operations to minimize the number of affected consumers after an incident.
Example 2.1.1 Consider the schematic power grid in Figure 2.2(i). The ring network is operated radially which, under normal operating conditions, is achieved by having an open switch to the right of transformer substation D. Now suppose an incident (e.g., a short circuit) occurs on the power supply line between substations C and D. The protection device (1) automatically clears the short circuit, thus leading to an interruption of supply for all substations on the corresponding branch (substations A, B, C, and D). The faulted line can be isolated by opening the switches (2a) and (2b), cf. Figures 2.2(ii) and (iii). The de-energized branch can be reconnected (3) and by closing the switch (4), all substations are supplied with power, cf. Figures 2.2(iii) and (iv).

![Figure 2.2](image)

Figure 2.2 Example of resupplying consumers by switching operations.

Low-voltage power grids are usually constructed and operated as radial networks. Medium-voltage power grids are usually constructed as (branched) rings but are normally operated as radial networks, i.e., the rings are opened at specific disconnection points. As a consequence, each
incident in MV/LV immediately leads to an interruption of supply. Due to the ring structure of MV grids, some of the affected consumers can usually be resupplied by switching operations (as was illustrated in Example 2.1.1). In general, there is no possibility of remote control in MV/LV and the switching operations have to be performed manually and on site. A resource thus has to travel to the location of the incident and work on site. In some cases, additional repair work by specialized resources is necessary to resupply the last customers. A temporary non-availability of resources, e.g., when all resources are working on other (more important) incidents, delays the restoration for the incident and thus directly influences the continuity of supply.

High- and extra-high-voltage power grids are built and operated with redundancies. The redundancies stem from the meshed construction of the grid and from the installation of parallel elements such as two parallel power supply lines or two transformers at the same location. The redundant elements are either connected to share the loading and absorb it if one element fails (so-called active or parallel redundancy) or remain on standby and are connected if the operating element fails (so-called standby redundancy).\(^{23}\)

In addition, high- and extra-high-voltage power grids possess a high degree of automation and allow the remote monitoring and control of these grids. Hence in general, a single incident in HV/EHV does not lead to an interruption of supply. HV/EHV grids thus satisfy the \((n - 1)\) criterion, which requires that a power grid can cope with the failure of any single element in the network, i.e., there is no interruption of supply, no expansion of the failure, and no damage of other pieces of equipment.\(^{24}\) Nevertheless, the failed equipment generally needs to be repaired or at least inspected and thus resources are required on site. Even though there is no immediate effect on the supply of consumers, an incident in HV/EHV reduces the redundancy in the power grid and additional incidents might eventually lead to a (large) interruption of supply. A temporary shortage of resources delays the repair and thus affects the duration of the grid being in a state of risk.

The restoration time for an incident (as defined in Section 2.1.2) might be considerably longer than the duration of the corresponding interruption of supply. An incident might not cause an interruption of supply e.g. due to the redundancies in the power grid or cause only a short interruption

\(^{23}\) cf. [BA92, p. 12].

\(^{24}\) Partly translated from [SBFN08, p. 40].
of supply because the affected customers can e.g. be resupplied by remote switching operations or by the use of emergency power generators. The durations of supply interruptions thus not only depend on the availability of resources for the restoration work but also on structural and technical aspects of the power grid.

In all voltage levels of the power grid, the availability of resources however has a strong impact on the restoration time for incidents and – as argued above – eventually on the continuity of supply. Hence, to take organizational decisions concerning resource management, it is crucial for power grid operators to analyze and quantify the effects of a limited availability of resources on the continuity of supply.

2.2 Modeling approach and two foci

In this work, we describe two models for analyzing the effects of a limited availability of resources on the continuity of supply. The models are based on discrete event simulation and Markov decision processes, respectively. Both models are state-based, where the state describes the state of the power grid and, if necessary, the location of the available resources for the restoration work after incidents. The state evolves over time and the subsequent state is determined by stochastic influences (e.g., random incidents of electrical components) as well as the chosen action/control of the power grid operator (e.g., repair of components by the resources). The models are designed to possess the “Markov property”, i.e., the choice of the action and the future development of the state only depend on the current state but not on the previously visited states and/or chosen actions. An important feature of both models is that the restoration times depend endogenously on the availability of resources. The goal is to make this implicit dependency explicit and thereby visible.

The problem of finding the “best” organizational structure and operational policy is split into two decision levels. On the first level, the organizational structure, i.e., the total number of resources (and, if necessary, their organization\textsuperscript{25}) in the system is specified. For a given organizational structure, the models are used on the second decision level, where the operational policy, i.e., the assignment of the available resources to the

\textsuperscript{25} e.g., the division of the supply area in areas of responsibility or the possibility of cross-boundary work, etc.
incidents, is decided. Being parameterized by the organizational structure, i.e., by the decision on the first level, the models can be applied in a hierarchical planning context. By repeatedly running the model (2nd level), the cost and the continuity of supply for different organizational structures (1st level) can be evaluated and compared. The two decision levels and their interaction are illustrated in Figure 2.3.

Figure 2.3 Illustration of the two decision levels in the modeling approach.

2.2.1 Two different foci

The two models we propose differ in their focus and the associated modeling “granularity”. The aim of the first model is to simulate the operational activities of the resources after incidents in the power grid. Incidents in all voltage levels (i.e., LV/MV/HV/EHV) are considered. Focusing on the organizational structure and the operational processes (e.g., assignment of resources), the physical processes within the power grid (e.g., the power flow) are neglected. To allow for a detailed evaluation, e.g., of delays in the restoration due to the traveling or a temporary unavailability of resources, many organizational and operational details are included, such as the incorporation of areas of responsibility and the distinction between periods of normal work and standby service, etc. On the other hand, the detailed modeling increases the model’s complexity and prohibits a direct optimization of the assignment policy (i.e., of the decisions on the second level). Due to its simulative nature, real-world instances of this model are computable.

By neglecting the processes within the power grid, the effects of simultaneous incidents cannot be assessed dynamically in the first model. Thus,
simultaneous incidents have an additive effect on the system. Whereas this assumption is usually realistic for incidents in medium- and low-voltage power grids, it might not adequately describe the effects in high- and extra-high-voltage. The second model tries to overcome this weakness by incorporating a model for the power grid, which assesses the effects of incidents with a DC power flow model. The focus of the second model lies on the “vulnerability” of power grids over time, in particular of power grids with redundancies (typically HV/EHV grids) where the effects of several incidents are superadditive. A single incident usually has no effect on the system (due to the \((n - 1)\) criterion) but the simultaneous occurrence of several incidents may lead to a large interruption of supply (blackout) and may thus have bad consequences for the system. Moreover, by simplifying the operational processes (in contrast to the first model), the second approach allows for an optimization of the decisions on the second level (i.e., of the assignment policy). Given a decision of the first level, the aim of the second model is to determine the (long-term) optimal assignment policy and the corresponding continuity of supply. The second model is formulated as a continuous-time Markov decision process. Even though such models allow for a direct optimization of the operational policy, the major difficulty in our setting lies in the very large number of states. To resolve this issue, we formulate an approximate model which is based on the aggregation of states.

In Table 2.1, the characteristics of the two models are summarized. The

<table>
<thead>
<tr>
<th>focus:</th>
<th>model/focus I</th>
<th>model/focus II</th>
</tr>
</thead>
<tbody>
<tr>
<td>voltage level:</td>
<td>LV/MV/HV/EHV</td>
<td>HV/EHV</td>
</tr>
<tr>
<td>modeling granularity:</td>
<td>many details</td>
<td>simplified</td>
</tr>
<tr>
<td>method:</td>
<td>discrete event simulation</td>
<td>Markov decision process</td>
</tr>
</tbody>
</table>

Table 2.1 Summary of the characteristics of the two models proposed in this work.
first model (focus I) is introduced in Chapter 3 and the second model (focus II) is presented in Chapter 5.

2.3 Related work

An important area of research for assessing the quality of power systems are reliability evaluations. The aim of reliability evaluations is to quantify the ability of a power system to supply power, e.g., by calculating reliability indices such as the average failure rate or the average annual outage time. Often, the components of the considered system are represented by state-based models and the reliability of the system is assessed either analytically or by simulation. Details on models and techniques for reliability evaluations are given in [BA92] and the sequel [BA96] that focuses on power systems. Typically, however, these evaluations do not take into account organizational characteristics (e.g., limited number of resources) of the power grid operator. Many papers suggest to use arbitrary (instead of traditionally exponential) probability distributions for the repair times of the components (e.g. [DS88], [Fon89]) and some authors mention that the restoration time depends on the travel time of the resources (e.g. [PTV04], [BVC04]) or the availability of resources (e.g. [ZSB08]). Nevertheless, the probability distributions of the repair/restoration times of the components are usually assumed to be independent of the availability of resources. [TT02] include travel and delay times of the resources in the restoration process which, however, are assumed to be constant.

There are a few works that do address the issue of a limited availability of resources. [WL07] analyze the effects of sequential switching operations and – if resources are available – parallel switching operations on reliability indices. In [WB02], restoration times that vary over time are introduced and the effect of the availability of resources is assessed by time-dependent factors that are used to adjust the restoration time. In contrast to [WB02], the restoration times in our work (focus I and II) are modeled endogenously and result from the inherent dynamics of the models. In the context of computer systems, [MdSeSG89] provide an approximation method for Markov chains to calculate upper and lower bounds on the availability of a (computer) system with limited resources for repair. Some of their ideas were used for (the analysis of) the aggregated model in Chapter 5.

Another major area of research, concerning the operation of power sys-
tems, is power system restoration. The aim of power system restoration research is to investigate how to re-establish the supply of power after a partial or a complete blackout. The main goal lies in finding sequences to start generating units and to reconnect power supply lines and loads. The focus lies on the technical aspects of the electrical components and the highly complex electrical processes in the power system. The restoration sequences are coordinated and controlled by operators in control centers. Organizational aspects of the power grid operator such as a limited number of resources are generally not considered. A commented collection of publications on power system restoration is provided in the book [Adi00]. A survey of papers is also given in [LDA01]. In [LVH+09], several new approaches for power system restoration are provided. The authors develop four modules that are intended to be integrated into an on-line decision support tool.

For the distribution system, a lot of research focuses on finding an optimal configuration of the power grid after an interruption of supply to resupply as much power (or as many customers) as quickly as possible. These models incorporate electrical constraints (e.g., capacity or voltage limits) as well as topological constraints to assure a radial structure of the resulting network. To implement a grid configuration, switching operations have to be performed. In practice, the switching operations often have to be performed manually in the field and thus the switching times strongly depend on the availability of human resources. In the survey article [COCL95], only 4 of the 19 reviewed papers consider the difference of remotely controlled and manually operated switching devices. However, these approaches as well as more recent work do not sufficiently consider the effects of a limited number of resources and their organization on the switching times. To reduce the restoration time, [FMH+92] and [COL97] give priority to switches that are controlled remotely over switches that are operated manually. [DM89] suggest an index that depends on the total switching time to classify restoration solutions, but the switching times are independent of the available resources. Many authors address the issue of manually operated switches and the associated switching times by reducing the number of locations to visit, i.e., by minimizing the number of switching operations (e.g. [WTC91], [HHK+92], [YH02]). In [ILD02], one part of the cost function to be minimized is the cost related to switching operations, which captures whether a switch is operated manually or controlled remotely, the time required to reach the switch, etc.
[RV05] analyze the effect of a limited availability of resources on the restoration time. The availability of resources is classified into three categories (few, regular, or sufficient), depending on how many incidents are simultaneously present. In our work (focus I and II), we improve this rather coarse-grained modeling of the availability of resources by considering each available resource individually.

[PAB+10a, PAB+10b] survey models and algorithms for the planning and operation of distribution systems. The authors group the corresponding problems according to their planning horizon into reliability planning problems (strategic and tactical level), cf. [PAB+10a], and contingency planning problems (operational level including real-time decisions), cf. [PAB+10b]. Strategic problems include capacity planning for substations, designing networks that are more reliable, partitioning of a supply area into service districts, or locating depots for resources. Only the latter two problems are related to the organization of a power grid operator. However, the corresponding models are usually based on an aggregated workload and do not consider the operational processes in detail. For example, [WSMT04] propose two models for locating depots. In the first model, locations of depots are determined such that the total transportation cost is minimized. Based on existing depots, the second model (which belongs to the tactical level) locates additional depots, which might be of particular interest if an extreme event with an increased workload occurs. [YM98] describe three organizational models with particular focus on situations with a heavy restoration workload due to extreme weather conditions. The tactical model determines how to redistribute the existing resources when extreme weather conditions are forecast. Using the tactical model as submodel, the short-term strategic model decides on the distribution of the existing resources to the depots under normal weather conditions. Based on the short-term strategic model, the long-term strategic model allows to acquire additional resources.

In order to support decisions on the organization of (human) resources with respect to the realizable continuity of supply, both the strategic and the operational level have to be considered. In particular and in contrast to the described approaches, more emphasis has to be put on the dependency of the restoration times on the availability of a limited number of resources.

A model for optimizing the organization of a power grid operator concerning both strategic and operational aspects is proposed in [ZDC92, ZDCD93]. The model aims at reducing the restoration times for in-
terruptions of supply by optimizing the organization and dispatching of emergency repair trucks. The authors suggest an iterative heuristic procedure to divide the service area (given by a set of “spatial units”) of a power grid operator into $N$ disjoint connected districts. Each district has a single emergency repair truck that is used for the restoration work associated with this district. Based on the average workload and the average repair time in each spatial unit, the districts are chosen such that the total travel time is minimized and such that workload and size of the districts are similar. The performance (e.g., average restoration time) in each district is assessed by simulation, where a set of incidents (characterized by time of occurrence, location, type, and priority) is generated and assigned to the (corresponding) emergency repair truck according to a priority-based first come first serve or a priority-based nearest neighbor dispatching rule. If the performance of the service area does not meet the target requirements, $N$ is increased to $N + 1$ and the model is solved again. In [ZDT98], the districting/simulation model is supplemented by a data-management and a vehicle-monitoring and communication module to form a decision support system for power grid operators. In addition to the described approach, our model (with focus I) also includes incidents without interruption of supply. To allocate resources to the incidents, we solve an assignment problem that maximizes the total “restoration efficiency” which is based on the interrupted power or power-at-risk (for incidents without interruption of supply) and the (estimated) restoration time for the incidents as well as the travel times of the resources.

[CD07] suggest a discrete event simulation model for power system restoration after an earthquake. The model includes many details of the real restoration process such as constraints on resources and equipment, different phases of restoration, etc. and represents uncertainties in travel and repair times, etc. In the repair phase, the resources are assigned to the components based on priorities that are derived from damage assessment (which is done in a previous phase). One main goal of the model is to estimate individual restoration curves, showing the amount of load that is restored within a certain amount of time, for different regions of the supply area. In contrast to the described model, we model the operations of a power grid operator in daily situations where new incidents keep occurring from time to time. As mentioned above, our model (with focus I) allocates the resources to the incidents based on an assignment problem instead of a priority-based rule. It is to mention that the model with focus I is also suited for extreme situations with an
increased number of incidents, e.g., due to a heavy storm, etc.

[ZSG+08] model the restoration process in a distribution system by simulating a queuing system with a limited number of resources. The failure and repair rates vary over time and the failures are assigned to the resources based on the first come first serve rule.
Chapter 3

Detailed grid operation model

(focus I)
In this chapter, we describe the first of our two models, with focus on the organization of the (human) resources of a power grid operator and the operational processes after incidents in the power grid (focus I, cf. Section 2.2.1). For a given organization of resources (1st decision level, cf. Section 2.2), the model simulates the resources’ activities for the restoration after incidents in all voltage levels. Aiming to mimic the real-world processes, many organizational and operational details are included. To decide how to allocate the available resources to the incidents (2nd decision level), an assignment problem is solved.

The model in this chapter was developed in a joint research project with RWE Rhein-Ruhr Netzservice GmbH, a German grid service company. The practical relevance of the project for RWE is emphasized in the press release [RWE07] by RWE Energy. Parts of this chapter are based on the joint papers [GFLZ08, GFB+09]. In this thesis, we focus the presentation on our main contribution, the mathematical modeling. The application of the model for the reorganization of a power grid operator requires a profound knowledge of the business environment. Indeed, the organizational implementation of the model’s results at RWE Rhein-Ruhr Netzservice GmbH has led to a complementary scientific work [Fri11]. [Fri11] provides an in-depth analysis and discussion of the model’s application for the reorganization of power grid operators and deduces a set of organizational criteria.

The chapter is organized as follows. The model is presented in Sections 3.1 – 3.5. Section 3.1 describes the restoration process of incidents in all voltage levels. The organization of the resources is characterized in Section 3.2. The randomness of the model is captured by scenarios, which is discussed in Section 3.3. The assignment of resources to incidents is presented in Section 3.4. A selection of key performance indicators for evaluating an organization of resources is provided in Section 3.5. In Section 3.6, the model is applied to a real-world supply area in a case study. We discuss how the obtained results can be used to evaluate and compare different organizations of resources and thus may support strategic organizational decisions.
3.1 Power grid, incidents, and restoration process

The focus of this model lies on the operational processes of a grid operator and in particular on the employment of (human) resources. The physical processes within the power grid (e.g., the power flow) are secondary. Consequently, we do not model the power grid with its physically connected pieces of electrical equipment as such. Instead, we divide the considered power grid into sufficiently small (geographically connected) areas and subsume all the electrical equipment of each such area into a so-called (aggregated grid) node (cf. the nodes in Figure 3.3 on page 33). An incident of an electrical component (e.g., a busbar, a transformer, a power supply line, etc.) is associated with an incident in the corresponding node and thus a demand for resources only occurs in the (aggregated grid) nodes. The second reason for aggregating the electrical components is to simplify the traveling process of the resources. In principle, however, the model allows to include one node for each electrical component. Each incident of an electrical component is characterized by a generic profile that represents the restoration process. The incidents are grouped into different types. These types of incidents allow the distinction between different voltage levels, types of affected electrical components (e.g., cable line, overhead line, etc.), grid design and state, etc. Due to the wide difference in the effects of a typical incident in MV/LV (with interruption of supply) and HV/EHV (without interruption of supply) and the specific characteristics of the corresponding power grids (cf. Section 2.1.3), the two cases are distinguished in the modeling of incidents. Whereas interruptions of supply can be characterized by the interrupted power, the effects on the continuity of supply are not immediately measurable for incidents without supply interruption. Therefore, a new measure, the so-called power-at-risk is introduced in Section 3.1.2. Throughout this chapter we assume that all incidents in MV/LV cause an interruption of supply.

In the following, we describe the modeling of the restoration process for incidents in MV/LV (Section 3.1.1) and incidents in HV/EHV (Section 3.1.2).
3.1.1 Incidents in MV/LV

In this section we describe the restoration process for an incident in MV/LV with an interruption of supply. We assume that resources which are working on an incident perform their tasks optimally, i.e., with the least possible amount of energy not delivered in time.

Let \( j \) denote an incident in MV with an interruption of supply and let \( v^j \) denote the aggregated grid node which comprises the corresponding electrical component. First we describe the development of the restoration process for \( j \) in the special case where all the required resources are already at the site of the interruption (i.e., at \( v^j \)) and are permanently available. Due to the assumption of optimally working resources, this special case yields the best possible (i.e., with a minimum amount of energy not delivered in time) restoration process. The corresponding generic profile of the restoration process is shown in the upper part of Figure 3.1. The time of occurrence of the incident is denoted by \( t_0^j \) and the initially interrupted power by \( P_{1j} \). During \( \delta_a^j \) the interruption has to be analyzed and no power can be restored. As MV power grids are generally constructed as branched rings, customers can usually be resupplied by switching operations (cf. Example 2.1.1 on page 14). Consequently, the interrupted power follows a decreasing step function (cf. dashed line in Figure 3.1(i)). To simplify the computations, this step function is approximated by a linear function. Hence, during \( \delta_s^j \), the interrupted power is assumed to follow a linearly decreasing function. It is assumed that there is (exactly) one type of resource which is qualified for the analysis/switching phase of all incidents in MV and that the analysis and the switching can be handled by a single resource. In some cases, additional component repair is required to resupply the last customers, which has to be carried out by specialized resources. Depending on the type of incident, the repair phase may require different types of resources. It is assumed that each type of incident requires a unique type of resource for the repair and that a single resource (of the correct type) can do the repair work. Of course, this single resource can also be interpreted as a team of resources. The repair phase is characterized by the (remaining) interrupted power \( P_{2j} \) and the repair time \( \delta_r^j \) needed by the specialized resources.

In the model, the characteristic data \( P_{1j}, \delta_a^j, \) and \( \delta_s^j \) of the incident are revealed as soon as the incident occurs and the need of an additional component repair, i.e., \( P_{2j}, \delta_r^j \), and the required type of resource, become
3.1 Power grid, incidents, and restoration process

Due to the assumption of optimally working resources, the working times \( \delta_a^j, \delta_s^j, \) and \( \delta_r^j \) as well as the development of the interrupted power in Figure 3.1(i) do not depend on the organization of the resources and thus are input data to the model.

Next we describe the (more realistic) general case where the required resources might not be available and/or have to travel to the site of the incident first. A corresponding possible development of the restoration process is shown in the lower part of Figure 3.1. As soon as possible after the occurrence of the incident \( t^j_0 \), a resource is assigned and sent to the node \( v^j \), the site of the interruption. At \( t^j_1 \), the resource arrives on site (for the first time). Between \( t^j_1 \) and \( t^j_2 \), the incident is analyzed (shaded area) or – if the resource is more urgently needed somewhere else – the analysis is postponed (blank area). Between \( t^j_2 \) and \( t^j_3 \), power is restored by switching operations (shaded area) or – again if the resource is called to another location – the restoration is postponed (blank area). As soon as possible after \( t^j_3 \), a suitable resource is assigned and at \( t^j_4 \), this resource arrives at the site of the incident (for the first time). Between \( t^j_4 \) and \( t^j_5 \), the repair work is carried out (shaded area) or once more postponed (blank area). As in the special case, the actual analysis requires \( \delta_a^j \) time.
units, the actual switching operations take $\delta^j_s$ time units, and the actual repair lasts for $\delta^j_r$ time units.

The travel times of the resources and the delay times (i.e., the times during which the restoration for the incident is postponed) strongly depend on the organization and availability of the resources. The model further includes two parameters for each type of incident that specify whether or not a resource which has started to work on an incident is allowed to leave before the work of the analysis/switching phase or the repair phase, respectively, is finished.

As in LV grids it is generally not possible to resupply customers by switching operations, it holds that $\delta^j_s = 0$ and $P^j_2 = P^j_1$ for interruptions in LV. In this case, the working time $\delta^j_a$ includes analysis and (the first part of the) restoration work and the corresponding phase is thus referred to as analysis/restoration phase. Other than that, the restoration process for an incident in LV (with interruption of supply) is identical to the one of an incident in MV. The resources for incidents in MV are also responsible for incidents in LV. In addition, each incident in LV has its specific number of customers that are affected by the interruption of supply. The number of affected customers in the analysis/restoration and repair phase are denoted by $N^j_1$ and $N^j_2$, respectively, and are revealed as soon as the interrupted power $P^j_1$ and $P^j_2$, respectively, are known. The values $N^j_1$ and $N^j_2$ are needed for the evaluation, cf. Section 3.5.

Summing up, an incident $j$ in MV is characterized by the input data

$$(t^j_0, v^j, \text{type}^j, P^j_1, \delta^j_a, \delta^j_s, P^j_2, \delta^j_r),$$

(3.1)

where type$^j$ denotes the type of the incident and where $P^j_2$ and $\delta^j_r$ are only revealed after the analysis/switching phase is finished. An incident $j$ in LV is specified by

$$(t^j_0, v^j, \text{type}^j, P^j_1, N^j_1, \delta^j_a, P^j_2, N^j_2, \delta^j_r),$$

(3.2)

where type$^j$ again denotes the type of the incident and where $P^j_2$, $N^j_2$, and $\delta^j_r$ become known only after the analysis/restoration phase is finished.

### 3.1.2 Incidents in HV/EHV

In the following, we describe the restoration process for an incident in HV/EHV without interruption of supply. During an incident without
interruption of supply, the redundancy in the power grid is reduced and additional incidents may lead to an interruption of supply. The extent of the (potential) supply interruption depends on where the additional incidents occur. The power-at-risk aims at quantifying this potential risk. For an incident without interruption of supply, let the (random variable) power-at-risk be defined as the potentially interrupted power in the event of additional incidents. The corresponding (empirical) probability distribution of the potentially interrupted power is determined by reliability evaluations. Namely, in a representative power grid, combinations of incidents are considered and the resulting interrupted power is determined. The combinations of incidents are categorized according to their types of incidents. For a given type of incident, the obtained values of the interrupted power then constitute the corresponding distribution of the power-at-risk by assuming that each value occurs with the same probability.

Let $j$ denote an incident in HV/EHV without interruption of supply and let $v^j$ denote the aggregated grid node which comprises the corresponding electrical component. As for incidents in MV, we begin with the special case where all the required resources are already at $v^j$ and are permanently available. Again assuming that resources which are working on an incident perform their tasks optimally (i.e., fastest possible in this case), this results in the best possible restoration. The corresponding generic profile is depicted in the upper part of Figure 3.2. The time of occurrence of the incident is again denoted by $t^j_0$. The power-at-risk of the incident, denoted by $P^j_{\text{risk}}$, is a sample drawn from the corresponding distribution of the power-at-risk. The obtained value of $P^j_{\text{risk}}$ is interpreted as the potentially interrupted power for the current state of the power grid. Due to the more complex technology in HV/EHV than in MV/LV, the different types of incidents in HV/EHV often require highly specialized resources. The incident is analyzed during $\delta^j_{a}$ by a single resource (of the correct type). It is assumed that each type of incident requires a unique type of resource. Depending on the type of incident, the restoration work might not be practicable for a single resource. Let $n^j_r \in \{1, 2\}$ denote the required number of resources for the restoration work. The restoration work time $\delta^j_{r}$ to clear the incident depends on the number of required resources $n^j_r$.

In the model, the characteristic data $P^j_{\text{risk}}$ and $\delta^j_{a}$ of the incident as well as the required qualification of the resource(s) is revealed as soon as the incident occurs and $n^j_r$ and $\delta^j_{r}$ become known only at the end of the
analysis. By the assumption of optimally working resources, the working times $\delta_a^j$ and $\delta_r^j$ as well as the number of required resources $n_r^j$ for the restoration work cannot be influenced and are input data to the model.

The lower part of Figure 3.2 shows a possible development of the restoration process in the (more realistic) general case where the required resources might be unavailable and/or have to travel to the site of the incident first. As soon as possible after the occurrence of the incident ($t_0^j$), a suitable resource is assigned and sent to $v^j$, the node of the incident. At $t_1^j$, the resource arrives on site (for the first time). Between $t_1^j$ and $t_2^j$, the incident is analyzed (shaded area) or – if the resource is called to another incident – the analysis is postponed (blank area). If necessary, as soon as possible after $t_2^j$, a suitable additional resource is assigned and this resource arrives on site (for the first time) at $t_3^j$. Between $t_3^j$ and $t_4^j$, the restoration work is carried out (shaded area) or once more delayed (blank area). As in the special case, the actual analysis requires $\delta_a^j$ time units and the actual repair lasts for $\delta_r^j$ time units. The model again includes two parameters for each type of incident that specify whether or not a resource which has started to work on an incident is allowed to leave before the work of the analysis phase or the restoration work phase, respectively, is finished.

The restoration process for an incident in HV/EHV with an interruption of supply is identical to the one of incidents without interruption of supply.
except that the incident is characterized by the interrupted power $P^j$ instead of the power-at-risk $P^j_{\text{risk}}$.

Summing up, an incident $j$ in HV/EHV without (or with) interruption of supply is characterized by the input data

\[
(t^j_0, v^j, \text{type}^j, P^j_{\text{risk}} (\text{or } P^j), \delta^j_a, n^j_r, \delta^j_r),
\]

(3.3)

where \( \text{type}^j \) again denotes the type of the incident and where \( n^j_r \) and \( \delta^j_r \) are only revealed at the end of the analysis phase.

### 3.2 Resources and their organization

For the restoration after incidents, resources need to travel to the corresponding (aggregated grid) nodes. Therefore, the nodes are connected by a set of edges (“roads”) and each edge is assigned an estimated (deterministic) travel time. The traveling of the resources is restricted to the graph defined by the (aggregated grid) nodes and the edges. Figure 3.3 shows an example of a graph (nodes and edges) of a supply area.\(^1\)

![Figure 3.3](image)

**Figure 3.3** Example of a graph of a supply area. The nodes of the graph are the aggregated grid nodes and the edges represent the “roads” on which the resources travel between the nodes.

Motivated by the restoration processes described in Sections 3.1.1 and 3.1.2, the model incorporates different types of specialists, each being qualified for certain tasks. For each type of specialist, a limited number of resources is allowed to travel on the aforementioned graph. These

---

\(^1\) Figure 3.3 was kindly provided by RWE Rhein-Ruhr Netzservice GmbH. It also appears in [GFB+09].
resources are organized in several areas of responsibility. An area of responsibility is a set of nodes where the associated resources are responsible for the restoration after incidents. The model incorporates the option of “cross-boundary work”, where resources are allowed to leave their (primary) area of responsibility to support the resources of another area. The availability of the resources and their areas of responsibility are time-dependent, allowing the distinction between periods of normal work (Monday to Friday from 07.30 until 16.30, i.e., daytime of working days) and standby service (which encompasses the remaining hours, i.e., nights and weekends). Whereas all resources are available during normal work, only a subset of (randomly chosen) resources is available during standby service. At the beginning of each period of normal work, the resources are uniformly distributed over the nodes of their corresponding areas of responsibility. This is to model that the resources are not waiting for the next incident to occur but in reality are occupied with maintenance work. At the beginning of each period of standby service, the selected resources are set to the node corresponding to their place of residence, where they remain on call in case of an incident.

Furthermore, the model allows a special form of collaboration of resources with different specializations. More precisely, the resources for the analysis/switching phase in MV and the analysis/restoration phase in LV are not only responsible for these phases of all incidents in MV/LV but they also analyze the incidents in HV/EHV. Then, depending on the type of incident, they either mobilize (and possibly support) the required specialized HV/EHV resource(s) or they are able to work on the incident on their own.

### 3.3 Randomness, gain of information, and scenarios

In the model, we consider a given time period \([0, T]\), e.g., one year. Let \(I\) denote the set of incidents in LV, MV, HV, and EHV that occur within this period and let \(R_t, t \in [0, T]\), denote the set of available resources at time \(t\).\(^{2}\) Due to the stochastic nature of incidents, \(I\) is a random object. As the time-dependent availability of the resources as well as

\(^{2}\) As the availability of resources maximally changes twice per day (at 07.30 and 16.30), \(R_t\) can fully be described by a finite set \(R_{tk}, k \in \{0, 1, \ldots, n\}\) for some \(n \in \mathbb{N}_0\).
their location at 07.30 from Monday to Friday is modeled by a random procedure (cf. Section 3.2), $R_t, t \in [0, T]$, are also random objects. Thus, there is an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which $I$ and $R_t$ are defined.

At a given point in time $t \in [0, T]$, only the incidents that have occurred so far are known, i.e., those $j \in I$ with $t^j_0 \leq t$. In addition, some of the attributes of an incident are only revealed during the restoration process (cf. Sections 3.1.1 and 3.1.2). Let $I_t \subseteq I$ denote the set of active incidents at time $t$, where an incident $j$ is called active at time $t$, if it has already occurred (i.e., $t^j_0 \leq t$) and if it has not been cleared yet. Thus, $I_t$ also depends on the actually implemented restoration process. Also, in the model, the availability of resources $R_t$ is only known up to $t$. We assume that there is a filtration $\mathcal{F}_t \subseteq \mathcal{F}, t \geq 0$, that models this successive gain of information (of incidents and their attributes as well as of the availability of resources) over time, i.e., $I_t$ and $R_t$ are assumed to be $\mathcal{F}_t$-measurable.

An input scenario for the model consists of realizations $I(\omega)$ of $I$ and $R_t(\omega)$ of $R_t, t \in [0, T]$, for some scenario $\omega \in \Omega$. We refer to $I(\omega)$ as scenario of incidents and to the collection $R_t(\omega), t \in [0, T]$, as scenario of availability. Of course, there is an associated collection of realizations $I_t(\omega)$ of $I_t, t \in [0, T]$.

### 3.3.1 Scenario of incidents

Let $I(\omega)$ denote a scenario of incidents. As $I(\omega)$ is a realization of a random object, the input data of each incident $j \in I(\omega)$ (cf. (3.1), (3.2), and (3.3)) themselves are realizations of random variables. For the ease of notation, we omit $\omega$ in the notation of the attributes.

By analyzing historical data, the frequencies of occurrence of the different types of incidents and the probability distributions of the input data $P^j_1, P^j_2, \delta^j_a, \delta^j_s, \delta^j_r, n^j_r$, etc., can be estimated for each type of incident. As it is not possible to derive the distributions of the power-at-risk $P^j_\text{risk}$ from historical data, they have to be determined by reliability evaluations in a preprocessing stage. The estimated distributions and frequencies of incidents are used to generate scenarios of incidents. By varying the distributions and the frequencies of incidents, more extreme scenarios (e.g., due to exceptionally adverse weather) can be generated.
3.4 Assignment of resources to incidents and “dynamics” of the model

Let $I(\omega)$ and $R_t(\omega)$, $t \in [0, T]$, denote an input scenario as defined in Section 3.3. At a given point in time $t$, the currently available resources $R_t(\omega)$ are assigned to the currently active incidents $I_t(\omega)$. According to the assignment, the resources travel to their assigned incidents and work on the failed equipment. The assignment at $t$ remains valid until either

- a traveling resource arrives at his destination,
- the current work phase of an incident is completed (i.e., analysis, switching, or repair for incidents in MV, analysis/restoration or repair for incidents in LV, and analysis or restoration for incidents in HV/EHV),
- the next incident becomes active, or
- the availability of resources changes due to a transition between periods of normal work and periods of standby service.

In any of the four cases, a new assignment is calculated for the new point in time $t'(> t)$. As we assume that for each phase of an incident, the required type of specialist is known, each type of specialist is considered separately.\textsuperscript{3}

3.4.1 Preliminary remarks and a special case

One objective of a grid operator is to maximize the continuity of supply over a given time period. This translates in our model to finding assignments of resources to incidents (at each point in time) such that the resulting continuity of supply in $[0, T]$ is maximized. At time $t$, however, only the subset $I_t(\omega) \subseteq I(\omega)$ of active incidents is known. The assignment decision at $t$ has to be taken based on this partial information and it is thus generally not possible to reach the global optimum. Moreover, as we discuss in the sequel, even if all information is known at the time of decision, it is hard to find the optimal assignment.

\textsuperscript{3} In the collaborating organization, if a resource for MV/LV is allowed to support a resource for HV/EHV in the restoration work, the work processes of these two types of specialists are not independent because the two types should work in parallel. However, for simplicity of the model, the two types are still considered separately.
3.4 Assignment of resources to incidents

We consider the following special case. We restrict ourselves to incidents in MV and assume that no incident requires an additional repair. As a consequence, all input data is known at the time of occurrence of the incident. It is assumed that all incidents $j \in I(\omega)$ occur at time $t^i_0 = 0$. In addition, we assume that once having started to work on an incident, the resource is not allowed to leave until the incident is fully cleared. The continuity of supply is indirectly measured by the *average system interruption duration index* (ASIDI)\(^4\) which is proportional to the total energy not delivered in time. Maximizing the continuity of supply thus translates to minimizing the total energy not delivered in time, which – under the given assumptions – is equivalent to minimizing

\[
\sum_{j \in I(\omega)} t^i_j P^j_1,
\]

where $t^i_j$ denotes the time a resource starts to work on incident $j$ (cf. Figure 3.1).

If there is only one resource available and the total working times are identical (i.e., $\delta^j_a + \delta^j_s =: \delta$ for all $j \in I(\omega)$ and some $\delta \in \mathbb{R}$)\(^5\), this problem is known as the *weighted traveling repairman problem*\(^6\) (with weights $P^j_1$), i.e., a repairman has to visit $n$ nodes, where the travel times between all nodes are given (i.e., the underlying graph can be assumed to be complete), such that the sum of the weighted waiting times of the nodes is minimized. If all weights are equal (i.e., $P^j_1 =: P$ for some $P \in \mathbb{R}$ for all $j \in I(\omega)$), the problem is called *traveling repairman problem*, *delivery man problem* or *minimum latency problem*\(^7\) and consists of finding a Hamiltonian cycle (in the complete graph) with a given starting node that minimizes the total (or average) waiting time of the nodes. The traveling repairman problem is well-known to be NP-hard.\(^8\) The generalization to $k$ resources (repairmen) is called *$k$-traveling repairmen problem*\(^9\).

The following, further restricted special case, gives the motivation for our approach of determining the assignment. In addition to the assumptions of the special case above, we assume that all incidents occur in the same node, i.e., no traveling of resources is involved. We also assume that there

---

\(^4\) cf. [IEE04]. The precise definition of ASIDI is given in Section 3.5.
\(^5\) Recall that we assume $\delta^j_t = 0$ for all incidents $j$.
\(^6\) cf. e.g. [GJT02].
\(^7\) cf. e.g. [GJT02].
\(^8\) cf. [SG76].
\(^9\) cf. [FHR07].
is only one resource available. Then, the problem of minimizing (3.4) can be solved in polynomial time, namely by sorting the ratios

\[ \frac{P^j}{\delta^j_a + \delta^j_s}, \quad j \in I(\omega), \]

in descending order and repairing the incidents in the corresponding order, as shown in the following theorem.

**Theorem 3.4.1** Let \( \sigma \) be a permutation of \( \{1, 2, \ldots, |I(\omega)|\} \) that specifies the order in which the restoration for the incidents in \( I(\omega) \) is performed.\(^{10}\) Then \( \sigma \) is an optimal order (i.e., leads to the minimum in (3.4)) for the restricted special case if and only if \( \sigma \) satisfies

\[ \frac{P^\sigma(j)}{\delta^\sigma_a(j) + \delta^\sigma_s(j)} \geq \frac{P^\sigma(j+1)}{\delta^\sigma_a(j+1) + \delta^\sigma_s(j+1)} \quad 1 \leq j \leq |I(\omega)| - 1. \]

The proof is provided in Appendix B.1 on page 190. The idea of the special cases and the proof of Theorem 3.4.1 are based on [Zen06].

In the above special case, the incidents are repaired in the order of decreasing “power (to be restored) per time (needed)” values. We use the same idea in the general assignment which is described next.

### 3.4.2 Assignment: general case

For a fixed type of specialist \( s \), let \( R^s_t(\omega) \) denote the corresponding set of available resources and \( I^s_t(\omega) \) the corresponding set of active incidents at time \( t \). Due to their direct influence on the continuity of supply, interruptions of supply are generally prioritized over incidents without supply interruption. However, if the size of the potential supply interruption (i.e., the power-at-risk) is considerably higher than the affected power of an interruption of supply, the resupply of the interruption might be postponed. To capture this relation and to compare incidents with and without interruption of supply, a *comparison factor* \( f_{\text{risk}} \) has to be specified such that 1 unit of interrupted power is considered equally important as \( f_{\text{risk}} \) units of power-at-risk. For an incident \( j \in I^s_t(\omega) \), define the *power* 

\(^{10}\) We identify each incident \( j \in I(\omega) \) with an integer in \( \{1, 2, \ldots, |I(\omega)|\} \).
equivalent
\[ \tilde{P}^j(t) := \begin{cases} P^j(t) & \text{if } j \text{ is an interruption of supply} \\ \frac{1}{P_{\text{risk}}} P^j & \text{otherwise (and for all } t) \end{cases} \]

where \( P^j(t) \) denotes the (remaining) interrupted power of incident \( j \) at time \( t \). Consequently, the power equivalent \( \tilde{P}^j(t) \) allows to prioritize incidents with and without supply interruption according to their (estimated) "severity".

We introduce the following binary variables. For each incident \( j \in I^s_t(\omega) \), let \( z_j \in \{0, 1\} \) with
\[ z_j = \begin{cases} 1 & \text{if incident } j \text{ is worked on} \\ 0 & \text{otherwise} \end{cases} \]
and for each resource \( i \in R^s_t(\omega) \) and each incident \( j \in I^s_t(\omega) \), let \( x_{ij} \in \{0, 1\} \) with
\[ x_{ij} = \begin{cases} 1 & \text{if resource } i \text{ is assigned to incident } j \\ 0 & \text{otherwise} \end{cases} \]

The efficiency \( w_{ij,t} \) of assigning resource \( i \in R^s_t(\omega) \) to incident \( j \in I^s_t(\omega) \) is defined as
\[ w_{ij,t} := \frac{\tilde{P}^j(t)}{d_{ij} + \tau_j(t)}, \]

where \( d_{ij} \) denotes the travel time of resource \( i \) to incident \( j \). For incidents in MV/LV, \( \tau_j(t) \) is the remaining duration of either \( \delta^j_a + \delta^j_s \) (during the analysis/switching or analysis/restoration phase) or \( \delta^j_r \) (during the repair phase). For incidents in HV/EHV, \( \tau_j(t) \) is either the remaining duration of \( \delta^j_a \) plus the expected value\(^\text{14}\) of \( \delta^j_r \) (during the analysis phase) or the remaining duration of \( \delta^j_r \) (during the restoration work phase). The value \( w_{ij,t} \) can be interpreted as marginal restoration efficiency as it indicates the amount of power equivalent that can be restored per time unit.

\(^{11}\) For the ease of notation, the dependency of the incidents’ input data on \( \omega \) is omitted.

\(^{12}\) Clearly, these variables depend on the scenario \( \omega \) and on the time \( t \). However, for the ease of notation, we omit the dependency on \( \omega \) and \( t \).

\(^{13}\) Once more, we omit the dependency on \( \omega \) for the ease of notation.

\(^{14}\) with respect to the required number of resources \( n^j \) for the restoration work.
efficiencies $w_{ij,t}$ are used to decide the priorities of the incidents to be processed in the event of a shortage of resources as well as to select the nearest resources in the event of a surplus of resources.

Recall that for incidents in HV/EHV, $n^j_r \in \{1, 2\}$ denotes the number of resources required for the restoration work (cf. Section 3.1.2). For an incident $j \in I^s_t(\omega)$, let $n^j_{r,t}$ denote the number of required resources for $j$ at time $t$. For incidents in HV/EHV, $n^j_{r,t} = 1$ as long as the incident has not been (fully) analyzed (i.e., $t < t^2_j$, cf. Figure 3.2) and $n^j_{r,t} = n^j_r$ if the analysis of the incident has been completed (i.e., $t \geq t^2_j$). For incidents in MV/LV, $n^j_{r,t} = 1$ for all $t$ (cf. Section 3.1.1).

Further, for an incident $j \in I^s_t(\omega)$, let $R^{s,\text{not}}_t(\omega, j) \subseteq R^s_t(\omega)$ denote the set of resources that are not allowed to work on incident $j$ (because the incident occurred outside their area of responsibility) and let $R^{s,\text{fixed}}_t(\omega, j) \subseteq R^s_t(\omega)$ denote the resources that have already been working on $j$ before $t$ and are not allowed to leave the incident. At the transitions between periods of normal work and periods of standby service, the availability of resources changes. Consequently, $R^{s,\text{fixed}}_t(\omega, j) := \emptyset$ for all incidents $j$ at these transitions.

The assignment of resources to incidents at time $t$ is calculated by the following optimization problem:

$$\max \quad \sum_{i \in R^s_t(\omega), j \in I^s_t(\omega)} w_{ij,t} \cdot x_{ij} \quad (3.5)$$

s.t. $$\sum_{j \in I^s_t(\omega)} x_{ij} \leq 1 \quad i \in R^s_t(\omega) \quad (3.6)$$

$$\sum_{i \in R^s_t(\omega)} x_{ij} = n^j_{r,t} \cdot z_j \quad j \in I^s_t(\omega) \quad (3.7)$$

$$x_{ij} = 0 \quad j \in I^s_t(\omega), i \in R^{s,\text{not}}_t(\omega, j) \quad (3.8)$$

$$x_{ij} = 1 \quad j \in I^s_t(\omega), i \in R^{s,\text{fixed}}_t(\omega, j) \quad (3.9)$$

$$x_{ij} \in \{0, 1\} \quad i \in R^s_t(\omega), j \in I^s_t(\omega) \quad (3.10)$$

$$z_j \in \{0, 1\} \quad j \in I^s_t(\omega).$$

The objective (3.5) is to maximize the total marginal restoration efficiency. The constraints (3.6) assure that each resource is assigned to at most one incident. The constraints (3.7) require that an incident $j$ which is worked on is assigned the required number of resources. The constraints (3.8) prohibit assignments that violate the areas of responsibility.
and (3.9) ensures that resources keep working on previously assigned incidents if they are not allowed to leave.

In the optimization problem, incidents with higher power equivalent \( \tilde{P}^j(t) \) are prioritized, e.g., interruptions in MV are mostly preferred to interruptions in LV. Also, given a single resource and two incidents with an equal power equivalent and equal working times, the resource is assigned to the closer of the two. In the event of a shortage of resources, the incidents with the least efficiencies will generally not be assigned and will be delayed.

The described problem of assigning the available resources to the currently active incidents belongs to the class of integer linear programming problems (ILP), which are well-known to be NP-hard in general. As the problem size in our setting is usually very small (few incidents and moderate number of resources), the problem is solved very quickly, e.g., with the branch-and-bound method.

**Remark 3.4.2**

(i) If the Simplex method is used to solve the relaxed linear programming problems in the branch-and-bound method, the integrality constraint (3.10) can be relaxed to \( 0 \leq x_{ij} \leq 1 \) (\( i \in R^*_t(\omega), j \in I^*_t(\omega) \)). The returned optimal solution has the same optimal objective function value and each \( x_{ij} \) is integral because whenever all \( z_j \in \{0, 1\} \) are fixed, the corresponding feasible set has only integral vertices.\(^{15}\)

(ii) If \( n_{j,t}^1 = 1 \) for all incidents \( j \in I^*_t(\omega) \), the constraint (3.7) can be replaced by \( \sum_{i \in R^*_t(\omega)} x_{ij} \leq 1 \) and the variables \( z_j \) can be omitted.\(^{16}\) In this case, the problem reduces to an assignment problem and (3.10) can be relaxed to \( 0 \leq x_{ij} \leq 1 \) (\( i \in R^*_t(\omega), j \in I^*_t(\omega) \)) without changing the optimal objective function value. Moreover, the feasible set of the resulting LP has only integral vertices,\(^{17}\) and the Simplex method returns an integral optimal solution.

---

\(^{15}\) The coefficient matrix of the inequality constraints (3.6) and the inequality constraints corresponding to (3.7) (replacing each equality constraint of (3.7) by two inequality constraints) is totally unimodular, cf. [Sch03, Theorem 5.21]. Thus, the claim follows from [Sch86, Theorem 19.1].

\(^{16}\) Replacing (3.7) by \( \sum_{i \in R^*_t(\omega)} x_{ij} \leq 1 \) in the original problem and adding the constraints \( z_j = \sum_{i \in R^*_t(\omega)} x_{ij}, j \in I^*_t(\omega) \), the feasible sets of the original and the modified problem are equal. Clearly, the variables \( z_j \) can be omitted in the modified problem.

\(^{17}\) Again, the corresponding coefficient matrix is totally unimodular (cf. [Sch03, Theorem 5.21]) and the claim follows from [Sch86, Theorem 19.1].
3.5 Evaluation and key performance indicators

After the restoration for all incidents $I(\omega)$ of an input scenario, key performance indicators can be calculated for this input scenario, e.g., indices for the continuity of supply or the total cost. As the input data depend on the random scenario $\omega$, the obtained indicators are realizations of random variables and their expected values or distributions have to be estimated. To this end, we run the model repeatedly for a finite number $m$ of (independent) scenarios $\omega_1, \ldots, \omega_m$. This can be interpreted as analyzing a given organization over $m$ succeeding years. To obtain a high accuracy of the estimated values, a large number $m$ of scenarios is desirable. However, for reasons of computation time, the chosen number of scenarios is usually small, e.g., $m = 20$. Due to the rather complex dependencies of the values to be estimated on the (random) input data, a detailed analysis of the quality of the estimates is hardly possible. It is important though to keep this accuracy issue in mind when interpreting the resulting values.

According to the regulatory or internal requirements, (a subset of) the estimated key performance indicators can be used to evaluate a given organization of resources. Moreover, by comparing the calculated key performance indicators of a (parameterized) set of different organizations of resources, the most adequate organization with respect to the given requirements can be determined. Examples of such comparisons are given in the case study in Section 3.6.

For a scenario of incidents $I(\omega)$, let $I^{LV}(\omega) \subseteq I(\omega)$ denote the set of all incidents in LV, $I^{MV}(\omega) \subseteq I(\omega)$ the set of all incidents in MV, $I^{(E)HV}(\omega) \subseteq I(\omega)$ the set of all incidents in HV/EHV with supply interruption, and $I^{(E)HV}_{\text{risk}}(\omega) \subseteq I(\omega)$ the set of all incidents in HV/EHV without supply interruption. More generally, let $\mathcal{T}$ denote a set of types of incidents, e.g., $\mathcal{T}$ could denote all incidents in LV and MV, and let $I^T(\omega) \subseteq I(\omega)$ denote the set of incidents whose type belongs to $\mathcal{T}$.

Among others, the following key performance indicators are estimated.

- Non-availability of supply, measured by the system average interruption duration index (SAIDI) for incidents in LV and by the average system interruption duration index (ASIDI) for incidents
3.5 Evaluation and key performance indicators

in MV, HV, or EHV.\textsuperscript{18} The non-availability of supply is an indirect measure for the continuity of supply. The lower the non-availability of supply, the better the continuity of supply.

For incidents $I^LV(\omega)$ in LV, the SAIDI is defined by (cf. Figure 3.1)

$$Q^L_U(\omega) := \frac{1}{N_{\text{total}}^{LV}} \sum_{j \in I^{LV}(\omega)} N_j \cdot (t_j^3 - t_j^0) + N_j \cdot (t_j^5 - t_j^3),$$

where $N_{\text{total}}^{LV}$ denotes the total number of customers served in LV.

For incidents $I^{MV}(\omega)$ in MV, the ASIDI is defined by

$$Q^M_U(\omega) := \frac{1}{L_{\text{total}}^{MV}} \sum_{j \in I^{MV}(\omega)} E^j,$$

where $L_{\text{total}}^{MV}$ denotes the total connected load in MV and $E^j$ the energy not delivered in time which is given by $E^j := \int_{t_0}^{t_5} P^j(t) \, dt$ (cf. Figure 3.1).

If the considered time period is 1 year, the unit of $Q^L_U(\omega)$ and $Q^M_U(\omega)$ is “time units per year” (usually “minutes per year”). $Q^L_U(\omega)$ and $Q^M_U(\omega)$ calculate the time during which each customer and each unit of connected load, respectively, were not supplied within the considered time period.

The ASIDI is defined analogously for the incidents $I^{(E)HV}(\omega)$ in HV/EHV (with supply interruption), with $L_{\text{total}}^{(E)HV}$, the total connected load in HV/EHV, instead of $L_{\text{total}}^{MV}$. $E^j$, the energy not delivered in time, is given by $E^j := P^j \cdot (t_j^4 - t_0^j)$ (cf. Figure 3.2).

We estimate the (expected value of the true) non-availability of supply by the sample mean, i.e., the average over all scenarios $\omega_1, \ldots, \omega_m$ of the corresponding scenario-dependent values. Thus we use

$$\hat{Q}^L_U := \frac{1}{m} \sum_{k=1}^{m} Q^L_U(\omega_k)$$

for LV and

$$\hat{Q}^M_U := \frac{1}{m} \sum_{k=1}^{m} Q^M_U(\omega_k)$$

\textsuperscript{18} cf. [IEE04].
In the case study, we show results of the total non-availability $\hat{Q}^{LV} + \hat{Q}^{MV}$. To measure the variability of the scenario-dependent values, we calculate the corresponding sample standard deviation $\hat{\sigma}^{LV+MV}$, where (assuming $m > 1$)

$$\hat{\sigma}^{LV+MV} := \sqrt{\frac{1}{m-1} \sum_{k=1}^{m} (Q^{LV}(\omega_k) + Q^{MV}(\omega_k) - \hat{Q}^{LV} - \hat{Q}^{MV})^2}.$$ 

- Duration until arrival on site $T^{T}_{\text{arrival}}$ for an incident whose type belongs to $\mathcal{T}$. This is the time from the occurrence of the incident until a resource arrives on site for the first time. This key performance indicator evaluates the organization’s ability to react to new incidents. It includes travel time as well as the time during which the restoration for the incident is postponed because no resource is available. The duration until arrival on site is an important internal performance measure but also a possible regulatory key figure (for incidents with an interruption of supply). For an incident $j \in I(\omega)$, the (realization of the) duration until arrival on site is given by $t^{j}_{1} - t^{j}_{0}$ (cf. Figures 3.1 and 3.2). Each incident $j \in \bigcup_{k=1}^{m} I^{T}(\omega_k)$ provides a realization of $T^{T}_{\text{arrival}}$.

In the case study, we show results of the empirical distribution function $\hat{F}^{T}_{\text{arrival}}(\cdot)$ of $T^{T}_{\text{arrival}}$, which is calculated as

$$\hat{F}^{T}_{\text{arrival}}(t) := \frac{1}{\sum_{k=1}^{m} |I^{T}(\omega_k)|} \sum_{k=1}^{m} \sum_{j \in I^{T}(\omega_k)} 1_{\{t^{j}_{1} - t^{j}_{0} \leq t\}} t \geq 0. \quad (3.11)$$

- Total restoration time $T^{T}_{\text{restore}}$ for an incident whose type belongs to $\mathcal{T}$. The total restoration time for an incident is the duration from the incident’s time of occurrence until its repair (cf. Section 2.1.2). In our model, full repair is required after incidents with an interruption of supply to (completely) re-establish the supply. Thus, in our model, the total restoration time for incidents with an interruption of supply is of particular interest if the regulator requires compensation payments for interruptions that last for more than an accepted limit. The (realization of the) total restoration time is given by $t^{j}_{5} - t^{j}_{0}$ for an incident $j \in I^{LV}(\omega) \cup I^{MV}(\omega)$ (cf. Figure 3.1)
and by $t^j_4 - t^j_0$ for an incident $j \in I^{(E)HV}(\omega) \cup I^{(E)HV}_{\text{risk}}(\omega)$ (cf. Figure 3.2). Each incident $j \in \bigcup_{k=1}^{m} I^{T}(\omega_k)$ provides a realization of $T_{\text{restore}}$. In the case study, we show results of the empirical distribution function $\hat{F}_{\text{restore}}(\cdot)$ of $T_{\text{restore}}$ which is calculated similarly to (3.11).

• Total cost. The total cost are calculated as the sum of labor costs, costs for restoration and compensation payments. The labor costs consist of labor costs for staff during normal work and standby service. The costs for restoration include costs for material, driving costs, and payments for restoration work during standby service. The compensation payments are payments to customers who suffered from supply interruptions longer than a given maximal standard.

3.5.1 New key performance indicators for incidents without interruption of supply

The calculation of e.g. ASIDI only includes incidents with an interruption of supply and hence most of the incidents in HV/EHV are not considered. In order to assess an organization of resources with respect to the performance in the HV/EHV grid, a new measure is needed. To evaluate incidents without interruption of supply, define the grid-at-risk duration (for a scenario of incidents $I(\omega)$) by

$$Q_{\text{risk}}(\omega) := \frac{1}{L^{(E)HV}_{\text{total}}} \sum_{j \in I^{(E)HV}_{\text{risk}}(\omega)} P_{\text{risk}}^j \cdot (t^j_4 - t^j_0),$$

where $L^{(E)HV}_{\text{total}}$ again denotes the total connected load in HV/EHV. $Q_{\text{risk}}(\omega)$ corresponds to the time duration during which each unit of the total connected load in HV/EHV was at risk (in this scenario of incidents). Again, we estimate the (expected value of the true) grid-at-risk duration by the sample mean, i.e., we calculate

$$\hat{Q}_{\text{risk}} := \frac{1}{m} \sum_{k=1}^{m} Q_{\text{risk}}(\omega_k).$$

$\hat{Q}_{\text{risk}}$ estimates the duration of the power grid being in a state of risk and can be interpreted as a reliability measure for the system. Although
having no (direct) influence on the continuity of supply, a too high value of $\hat{Q}_{\text{risk}}$ is undesirable as it indicates an increased risk of additional interruptions of supply. We also calculate the corresponding sample standard deviation (assuming $m > 1$)

$$
\hat{\sigma}_{\text{risk}} := \sqrt{\frac{1}{m-1} \sum_{k=1}^{m} (Q_{\text{risk}}(\omega_k) - \hat{Q}_{\text{risk}})^2}.
$$

The grid-at-risk duration $Q_{\text{risk}}(\omega)$ is an integrated index, i.e., it reflects a whole set of incidents and thus obviously, the effect of a single incident cannot be extracted. However, it might be interesting to analyze which amount of power was effectively at risk for how long. To this end, we introduce the sorted power-at-risk duration curve.\textsuperscript{19,20} For a given time period, the sorted power-at-risk duration curve shows the relative time duration (on the horizontal axis) during which at least the corresponding amount of power (on the vertical axis) was at risk.

The sorted power-at-risk duration curve can be derived from the chronological development of the total power-at-risk at each point in time $t$ in the considered time period $[0, T]$ by sorting the blocks “power-at-risk versus time duration” with decreasing power-at-risk values. To estimate the (true) sorted power-at-risk duration curve, we consider the time horizon $[0, mT]$ and assume that the incidents $I_{\text{risk}}^{(E)HV}(\omega_k)$, $1 \leq k \leq m$, occur in $[(k-1)T, kT]$, i.e., the estimation of the sorted power-at-risk duration curve is based on $m$ subsequent periods (e.g., years). The sorted power-at-risk duration curve is illustrated in the following example.

**Example 3.5.1** The chronological development of the total power-at-risk is given in the upper part of Figure 3.4. The time horizon is chosen to be 16 hours. The different values of the total power at risk and the corresponding duration (absolute and relative with respect to the considered time horizon) are given in the following table. The corresponding sorted power-at-risk duration curve is given in the lower part of Figure 3.4.

\textsuperscript{19} Motivated by the load duration curve, cf. [BA96, p. 38].

\textsuperscript{20} The sorted power-at-risk duration curve is also analyzed in [Mos08].
3.5 Evaluation and key performance indicators

Figure 3.4 Example of the sorted power-at-risk duration curve.

<table>
<thead>
<tr>
<th>(total) power-at-risk</th>
<th>duration [h]</th>
<th>relative duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>2</td>
<td>12.5%</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>6.25%</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>12.5%</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>18.75%</td>
</tr>
<tr>
<td>0</td>
<td>8</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>100%</td>
</tr>
</tbody>
</table>
3.6 Case study

This section presents several possible evaluations of the model applied to a real supply area. The aim is to illustrate how the relation between an organization of resources and the desired key performance indicators can be analyzed qualitatively and quantitatively.

In order to apply the model in a realistic setting, a large amount of data had to be collected, processed, and analyzed. This intricate task was performed by RWE Rhein-Ruhr Netzservice GmbH (RWE). The whole set-up of the case study and the input data were kindly provided by RWE, which we greatly acknowledge. In particular, the design of meaningful organizational variants was also done by RWE. As mentioned, the application of the model in a real business environment and the corresponding organizational implications are analyzed and discussed in [Fri11]. The case study reported here and the interpretations of the results are mainly based on [GFLZ08, GFB+09].

The case study consists of two parts:

(i) MV/LV grid. In Section 3.6.1, a real MV/LV grid and only incidents with an interruption of supply are considered.

(ii) Combination of HV/EHV and MV/LV grid. In Section 3.6.2, a real EHV/HV/MV/LV grid is considered. Results of solely the HV/EHV grid and incidents without interruption of supply are presented as well as results of a combined model with incidents in the HV/EHV grid as well as in the underlying MV/LV grid and partly collaborating resources.

A summary of the case study is provided in Section 3.6.3.

3.6.1 MV/LV grid

The area of the considered MV/LV grid in this section has an approximate size of 2500 km$^2$ and includes both rural and urban zones. The corresponding grid consists of approximately 13,500 km of cable lines (MV+LV), 6100 km of overhead lines (MV+LV), and 6000 transformer substations MV/LV. In the model, this area is represented by 55 nodes, each covering an area with a diameter of approximately 8 km. The nodes are connected by 179 edges with an average travel time of 21 minutes.

In this part of the case study, two fixed scenarios of incidents (based on
historical data) are used. Table 3.1 shows the key data of the scenarios of incidents. Scenario “normal” corresponds to a typical set of interruptions during one year and has a rate of 1 interruption per 8.2 hours. Scenario “extreme” models a short period (58 hours) with a large number of interruptions, e.g., due to a storm, and has a rate of 1 interruption per 16 minutes.

<table>
<thead>
<tr>
<th>scenario</th>
<th>number of interruptions</th>
<th>period</th>
</tr>
</thead>
<tbody>
<tr>
<td>“normal”</td>
<td>203</td>
<td>866</td>
</tr>
<tr>
<td>“extreme”</td>
<td>99</td>
<td>120</td>
</tr>
</tbody>
</table>

Table 3.1 Scenarios of incidents.

In the considered scenarios of incidents, no incident requires an additional repair, i.e., $\delta^j_r = 0$ for all incidents $j$ (cf. Figure 3.1). As a consequence, only one type of resources is required. An organization of these resources is characterized by four parameters, which are defined as

- $\rho_N$ the relative number of resources during periods of normal work given as ratio to the current (as of the year 2006) absolute number ($= n_N^0$), e.g., $\rho_N = 60\%$;
- $n_A$ the number of areas of responsibility;
- $n_S$ the (absolute) number of resources during periods of standby service, and
- $1_A \in \{\text{yes, no}\}$, indicating whether (“yes”) or not (“no”) the resources are allowed to work outside their area of responsibility (cross-boundary work).

Five different organizations of resources are considered, each with a varying number $\rho_N$ of resources during periods of normal work. The parameters $n_A$, $n_S$, and $1_A$ of the organizations are given in Table 3.2. Whereas all organizations $O_k$, $k \neq 3$, have a fixed number of resources during standby service for all values of $\rho_N$, the corresponding value $n_S$ for organization $O_3$ is equal to $\rho_N \cdot n_N^0$, the number of resources during normal work. Hence, the resources in organization $O_3$ are permanently available.

---

21 To better capture the randomness of incidents and obtain more reliable results, the scenarios of incidents would have to be generated randomly.

22 For reasons of confidentiality, the value of $n_N^0$ is not provided.

23 For each value of $\rho_N$ and for each input scenario, the resources’ places of residence are chosen randomly.
Table 3.2 Parameterized organizations of resources.

<table>
<thead>
<tr>
<th>organization (MV/LV)</th>
<th>$n_A$</th>
<th>$n_S$</th>
<th>$\lambda_A$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O_1$</td>
<td>11</td>
<td>11</td>
<td>no</td>
</tr>
<tr>
<td>$O_2$</td>
<td>11</td>
<td>11</td>
<td>yes</td>
</tr>
<tr>
<td>$O_3$</td>
<td>11</td>
<td>$\rho_N \cdot n_N^0$</td>
<td>yes</td>
</tr>
<tr>
<td>$O_4$</td>
<td>8</td>
<td>8</td>
<td>yes</td>
</tr>
<tr>
<td>$O_5$</td>
<td>4</td>
<td>4</td>
<td>yes</td>
</tr>
</tbody>
</table>

Evaluations for the following key performance indicators are presented:

- non-availability of supply,
- duration until arrival on site, and
- total restoration time for an interruption.

In this part of the case study, 20 input scenarios were used. As the two scenarios of incidents are fixed, the 20 input scenarios are only needed to average the random effects of the availability of the resources.

**Non-availability of supply**

Figure 3.5(a) shows the (estimated) total non-availability of supply $\hat{Q}_U^{LV} + \hat{Q}_U^{MV}$ (cf. Section 3.5) for the “normal” scenario of incidents and all five organizations against different numbers of available resources during normal work. The upper (lower) endpoints of the vertical bars indicate the sample mean plus (minus) the sample standard deviation $\hat{\sigma}_U^{LV+MV}$ and are intended to give an idea of the reliability of the results. The variability of the results stems from the random effects related to the availability of the resources. For instance, the resources’ places of residence are randomly chosen in each input scenario and the resources that are working during periods of standby service are randomly selected at the beginning of each such period. The only exception is $O_3$. As all resources are permanently available, there are no random effects for this organizational variant and thus the corresponding sample standard deviations are 0.

For all organizations except $O_3$, the number of resources during normal work has (almost) no significant influence on the total non-availability of supply. First, it is to remark that the current number of resources ($\rho_N = 100\%$) is not laid out for the restoration after interruptions only but also for maintenance work, which amounts for about 90% of the
yearly volume of work.\textsuperscript{24} Secondly, due to the small number of simultaneous interruptions and the prioritization of MV-interruptions in the assignment, almost no MV-interruption has to be delayed, even if there are only few resources. As a consequence, only an increasing number of LV-interruptions has to be delayed with fewer resources. As the contribution of LV to the total non-availability of supply is considerably smaller than the one of MV, this key performance indicator is hardly influenced by the considered number of resources. The exception is $O_3$, where all resources are permanently available. During periods of standby service, $O_3$ thus profits from a much larger number of available resources than the other organizations.

On the other hand, there are differences between the non-availability of supply of the organizational variants. Due to the permanent availability of the resources, organization $O_3$ outperforms the other organizations and provides a lower bound for the total non-availability of supply. Organization $O_2$ performs better than organization $O_1$ for all considered numbers of resources. This is due to the fact that the resources of organization $O_2$ are allowed to work outside their area of responsibility. Organization $O_4$ results in approximately the same total non-availability of supply as organization $O_1$. The effect of having fewer resources during standby service is compensated by allowing the resources to work outside their area of responsibility. In addition, organization $O_2$ outperforms $O_4$. It can be concluded that the number of resources during standby service, which is larger for organization $O_2$, leads to this effect. Finally, a further reduc-

\textsuperscript{24} cf. [Küp05, p. 3].
tion of the number of resources during standby service as in $O_5$ results in a significantly higher total non-availability of supply.

In Figure 3.5(b), the (estimated) total non-availability of supply $\hat{Q}^{LV}_U + \hat{Q}^{MV}_U$ for the “extreme” scenario of incidents is depicted for all five organizations. Again, the sample standard deviation $\hat{\sigma}^{LV+MV}_U$ is indicated with vertical bars. On the displayed scale, the vertical bars are hardly visible which suggests a good reliability of the results. The variability of the results is significantly smaller than in Figure 3.5(a). A possible explanation is that the considered time horizon in the “extreme” scenario is much smaller and thus there are less random effects related to the availability of the resources.

Contrary to the “normal” scenario, the total non-availability of supply reacts more sensitively to the number of resources during normal work. This is due to the large number of simultaneous interruptions in the “extreme” scenario. Having all resources permanently available, organization $O_3$ again outperforms the other organizations and provides a lower bound on the non-availability of supply. Organizations $O_4$ and $O_5$ perform considerably better than organization $O_1$. Allowing the resources to work outside their area of responsibility thus strongly improves the non-availability of supply. This is also confirmed by the significant difference between $O_1$ and $O_2$.

**Duration until arrival on site**

Figure 3.6(a) shows the empirical distribution function of the duration until arrival on site $T^{LV,MV}_{\text{arrival}}$ of all incidents (LV and MV) of the “normal” scenario of incidents for organization $O_2$ with $\rho_N \in \{14\%, 57\%\}$ and organization $O_5$ with $\rho_N \in \{16\%, 57\%\}$. Whereas for both variants of $O_2$, more than 90% of all interruptions have a duration until arrival on site of at most 30 minutes, only 73% of the interruptions achieve this value for organization $O_5$ with $\rho_N = 57\%$ (and 68% with $\rho_N = 16\%$). Again, the results indicate a strong influence of the number of resources during standby service. As organization $O_5$ has much less resources during standby service than $O_2$, more interruptions are delayed and the duration until arrival on site is worse.

For the “normal” scenario of incidents, the duration until arrival on site has a very similar distribution for interruptions in LV and interruptions in MV. This is illustrated for organizations $O_2$ (with $\rho_N = 14\%$) and $O_5$
(with $\rho_N = 16\%$) in Figure 3.6(b). However, for the “extreme” scenario of incidents, these distributions can vary considerably, as is depicted in Figure 3.6(c) for organization $O_2$ (with $\rho_N = 57\%$). In the “extreme” scenario, many interruptions are simultaneously present and as the assignment prioritizes interruptions in MV, the interruptions in LV are delayed much longer.

Observe that for all organizations there is a positive probability of a duration until arrival on site of 0. This effect is due to the aggregation of the electrical equipment into (aggregated grid) nodes in the model. When an incident occurs, a resource might already be at the corresponding node and thus can immediately start to work. Travel times within aggregated grid nodes, which are not captured by the model but do occur in reality, would have to be added to obtain more realistic results.
Total restoration time for an interruption

In Figures 3.7(a) and (b), the empirical distribution functions of the total restoration times $T_{\text{MV}}^{\text{restore}}$ and $T_{\text{LV}}^{\text{restore}}$ for all MV- and all LV-interruptions, respectively, of the “extreme” scenario of incidents are depicted for organization $O_2$ with $\rho_N \in \{14\%, 57\%\}$ and organization $O_5$ with $\rho_N \in \{16\%, 57\%\}$. Recall that in our model, the total restoration time for an incident equals the (maximum) duration of the corresponding interruption of supply. For illustration purposes, we assume that a compensation payment is due for interruptions that last for more than 12 hours. The vertical lines in the figures indicate a restoration time of 12 hours.

In the “extreme” scenario in MV (cf. Figure 3.7(a)), slightly less than 20% of the interruptions exceed the threshold value for both variants.
of $O_2$ and approximately 31% of the interruptions for the variants of $O_5$. In the “extreme” scenario in LV (cf. Figure 3.7(b)), not only many more interruptions exceed the threshold but the difference for different numbers of resources during normal work is also much larger. For $O_2$, approximately 65% and 79% of the interruptions exceed the threshold of 12 hours for $\rho_N = 57\%$ and $\rho_N = 14\%$, respectively. For $O_5$, around 83% and 94% of the interruptions exceed the threshold of 12 hours for $\rho_N = 57\%$ and $\rho_N = 16\%$, respectively. Furthermore, Figure 3.7(b) shows that LV-interruptions cannot be neglected in the event of an extreme scenario, although their contribution to the highly aggregated key performance indicator non-availability of supply might be marginal. Beside the compensation payments, such long-lasting LV-interruptions lead to a bad reputation of the power grid operator.

In the “normal” scenario of incidents, only very few interruptions exceed the threshold of 12 hours, as is illustrated in Figure 3.7(c). The figure shows the empirical distribution function of the total restoration time $T_{\text{restore}}^{\text{LV, MV}}$ for all MV- and all LV-interruptions for organization $O_2$ with $\rho_N \in \{14\%, 57\%\}$ and organization $O_5$ with $\rho_N \in \{16\%, 57\%\}$.

### 3.6.2 Combination of HV/EHV and MV/LV grid

As mentioned in Section 3.2, the model allows for organizations where the resources for MV/LV support the resources for HV/EHV for specific tasks. In this section, we compare such a collaborating organization to organizational variants with completely separated resources for MV/LV and HV/EHV.

The area of the power grid considered in this section has an approximate size of 8350 km$^2$ and again covers both rural and urban zones. The corresponding HV/EHV grid consists of approximately 5100 km of overhead lines and 240 transformer substations with 440 transformers. The underlying MV/LV grid encompasses approximately 21’000 km of cable lines, 7500 km of overhead lines, and 12’000 transformer substations MV/LV. The area is modeled by 322 nodes, each representing an area with an approximate diameter of 6 km. The nodes are connected by 978 edges with an average travel time of 21 minutes.

We consider a fixed scenario of incidents (based on historical data). More precisely, we fix all the input data except for the power-at-risk values and the required number of resources of the incidents in HV/EHV as these
values are not available from historical data. The scenario of incidents used in this part of the case study consists of 2053 incidents in HV/EHV, 299 incidents in MV, and 1257 incidents in LV. The scenario corresponds to a typical set of incidents in one year and has a rate of 1 incident per 4.3 hours in HV/EHV and 1 incident per 5.6 hours in MV/LV. Whereas all incidents in MV/LV result in an interruption of supply, all incidents in HV/EHV are without interruption of supply. Again, no incident in MV/LV requires an additional repair, and thus only one type of resources is required for the restoration in MV/LV. The organization of each type of specialist in HV/EHV is characterized by the four parameters \( (\rho_N, n_A, n_S, 1_A) \) defined in Section 3.6.1. In the following, the organization of all resources of the different qualifications in HV/EHV is represented by two aggregated parameters, namely

\[
\hat{\rho}_N \quad \text{the relative total number of resources (summed over all qualifications in HV/EHV) during periods of normal work given as ratio to the current (as of the year 2007) absolute number (=: \hat{n}_N^0)}^{25}, \text{ and}
\]

\[
\hat{n}_S \quad \text{the total (absolute) number of resources (summed over all qualifications in HV/EHV) during periods of standby service.}
\]

The resources of all qualifications are allowed to work outside their area of responsibility (i.e., \( 1_A = \text{“yes”} \) for each qualification). Two different organizations of resources in HV/EHV are considered, each with a varying number \( \hat{\rho}_N \) of resources during periods of normal work.\(^{26} \) The values of \( \hat{n}_S \) are given in Table 3.3. One organizational variant with collaborating resources in HV/EHV and MV/LV is considered. In this organization, the resources of organization \( \hat{O}_2 \) are supported by the resources of the underlying MV/LV grid. The corresponding organization of resources in MV/LV is denoted by \( O_6 \) and has \( n_A = 14 \) areas of responsibility and \( n_S = 15 \) resources during periods of standby service. \( O_6 \) is also considered with a varying number \( \rho_N \) of resources during periods of normal work and the resources are allowed to work outside their area of responsibility (i.e., \( 1_A = \text{“yes”} \)). The collaborating organization is denoted by \( \hat{O}_2 + O_6 \). In \( \hat{O}_2 + O_6 \), to account for differences in the qualification of the resources in MV/LV compared to the resources in HV/EHV, the working times of the incidents in HV/EHV as well as the probability distributions of the

---

\(^{25} \) Again, for reasons of confidentiality, the value of \( \hat{n}_N^0 \) is not revealed.

\(^{26} \) Again, the resources’ places of residence are chosen randomly for each value of \( \rho_N \) and for each input scenario.
number of required resources for the restoration work of these incidents were modified.

Evaluations for the following key performance indicators are presented:

- grid-at-risk duration,
- non-availability of supply for MV/LV,
- duration until arrival on site, and
- sorted power-at-risk duration curve.

In this part of the case study, 25 input scenarios were used. In this case, the input scenarios are needed to average the effects of the randomly generated power-at-risk values and number of required resources of the incidents in HV/EHV as well as the random effects of the availability of the resources.

**Grid-at-risk duration and non-availability of supply**

Figure 3.8 shows the (estimated) grid-at-risk duration $\hat{Q}_{\text{risk}}$ (cf. Section 3.5.1) for the organizations $\hat{O}_1$, $\hat{O}_2$, and $\hat{O}_2 + O_0$ against different numbers of available resources in HV/EHV during normal work. Again, the upper (lower) endpoints of the vertical bars indicate the sample mean plus (minus) the sample standard deviation $\hat{\sigma}_{\text{risk}}$ and shall give an idea of the reliability of the results.

For all organizations, there is (almost) no significant change in the grid-at-risk duration except for large differences in the number of resources. As in MV/LV, the current number of resources ($\hat{\rho}_N = 100\%$) is not only responsible for the restoration after incidents but also processes a large amount of maintenance work.

However, there are differences between the different organizational variants. The grid-at-risk duration for organization $\hat{O}_2$ is significantly worse than that of $\hat{O}_1$, which is explained by the fact that $\hat{O}_2$ has only about half as many resources during standby service as $\hat{O}_1$. This increase in the grid-at-risk duration (i.e., decrease in the reliability of the system) can

<table>
<thead>
<tr>
<th>organization (HV/EHV)</th>
<th>$\hat{n}_S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{O}_1$</td>
<td>19</td>
</tr>
<tr>
<td>$\hat{O}_2$</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 3.3 Total number of resources during periods of standby service.
be partly compensated by the collaborating organization \(\hat{O}_2 + O_6\). This compensation is even achieved without a significant decrease in the non-availability of supply. This is illustrated in Figure 3.9, which shows the (estimated) total non-availability of supply \(\hat{Q}_U^{LV} + \hat{Q}_U^{HV}\) for the MV/LV organization \(O_6\) with and without supporting the HV/EHV organization \(\hat{O}_2\). These results indicate the large advantage of a collaboration between resources in HV/EHV and MV/LV. Even a significant reduction in the number of resources during standby service in HV/EHV only leads to a moderate decrease in the reliability of the system (measured by the grid-at-risk duration) and, most importantly, the non-availability of supply is not significantly affected.

\[\text{Figure 3.8} \quad \text{Grid-at-risk duration } \hat{Q}_{\text{risk}} \text{ for organizations } \hat{O}_1, \hat{O}_2, \text{ and } \hat{O}_2 + O_6.\]

\[\text{Figure 3.9} \quad \text{(Estimated) total non-availability of supply } \hat{Q}_U^{LV} + \hat{Q}_U^{HV} \text{ for organization } O_6 \text{ with and without supporting the HV/EHV organization } \hat{O}_2.\]
3.6 Case study

Duration until arrival on site

Figure 3.10 shows the empirical distribution function of the duration until arrival on site $T_{\text{arrival}}^{(E)HV,\text{risk}}$ for all incidents in HV/EHV without interruption of supply for the organizations $\hat{O}_2$ and $\hat{O}_2 + O_6$ and for $\hat{\rho}_N = 4.7\%$ and $\hat{\rho}_N = 71\%$. The results again show that the collaborating organization is clearly beneficial. For the collaborating organization $\hat{O}_2 + O_6$, approximately 94\% (for $\hat{\rho}_N = 71\%$) and 90\% (for $\hat{\rho}_N = 4.7\%$) of all incidents have a duration until arrival on site of at most 50 minutes. On the other hand, for $\hat{O}_2$, only about 64\% (for $\hat{\rho}_N = 71\%$) and only 32\% (for $\hat{\rho}_N = 4.7\%$) of the incidents are reached within this time.

![Figure 3.10](image-url)

Figure 3.10  Empirical distribution functions of the duration until arrival on site $T_{\text{arrival}}^{(E)HV,\text{risk}}$ for all incidents in HV/EHV without interruption of supply for organizations $\hat{O}_2$ and $\hat{O}_2 + O_6$ (with $\hat{\rho}_N \in \{4.7\%, 71\%\}$).

Sorted power-at-risk duration curve

Figure 3.11 shows (a part of) the sorted power-at-risk duration curve (as defined in Section 3.5.1) for organizations $\hat{O}_2$ and $\hat{O}_2 + O_6$ and for $\hat{\rho}_N = 4.7\%$ and $\hat{\rho}_N = 71\%$. For all organizational variants, the sorted power-at-risk duration curve decays very fast. For instance, in less than 10\% of the time, more than 250 MW of power are at risk. In the shown domain, organization $\hat{O}_2 + O_6$ outperforms $\hat{O}_2$ for $\hat{\rho}_N \in \{4.7\%, 71\%\}$, i.e., the values of the power at risk duration curve of $\hat{O}_2 + O_6$ are strictly smaller than those of $\hat{O}_2$. Again, the curves confirm the advantage of the collaborating organization.
Figure 3.11 Sorted power-at-risk duration curve for organizations $\hat{O}_2$ and $\hat{O}_2 + O_6$ with $\hat{\rho}_N \in \{4.7\%, 71\%\}$.

3.6.3 Summary

The case study illustrates how the calculated key performance indicators can be used to evaluate and compare different organizations of resources. With the new key performance indicators that are based on incidents without interruption of supply, the performance of organizations of resources for HV/EHV power grids, where most of the incidents are without interruption of supply, can be assessed. Organizations which allow the resources to work outside their area of responsibility perform better than organizations which do not. The results show that effects of having fewer resources during standby service may be compensated by allowing the resources to work outside their areas of responsibility. Further, a collaborating organization where the resources of MV/LV support the resources of HV/EHV proves to be beneficial.

It is to remark that not all key performance indicators react equally sensitive to variations in the organization. For instance, in the “extreme” scenario of incidents, even though the non-availability of supply of organization $O_2$ does not vary much for different numbers of resources (cf. Figure 3.5(b)), the total restoration time for interruptions in LV changes significantly (cf. Figure 3.7(b)). Organizational decisions should thus be based on a whole set of indicators.

Our industrial partner, RWE Rhein-Ruhr Netzserservice GmbH, with whom the model in this chapter was developed, successfully applied this model in their reorganizational process. The model and results were used to evaluate different organizational variants and support strategic decisions.
Chapter 4

On Markov chains and Markov decision processes
In this chapter we review the theory of Markov chains and Markov decision processes. The model with focus II, which is described in Chapter 5, is formulated as a continuous-time Markov decision process. As the “backbone” of Markov decision processes are Markov chains, we first review their definitions and properties in Section 4.1. Then, in Section 4.2, we summarize the theory and algorithms of Markov decision processes with the average expected cost criterion. One key property in view of the application in Chapter 5 is that continuous-time Markov decision processes can be reformulated as discrete-time Markov decision processes.

4.1 On homogeneous Markov chains

Consider a system whose state evolves over time, due to some random influences. The set of possible states is denoted by $S$ and is referred to as state space. The general theory of Markov chains usually considers countable state spaces. However, we limit our discussion to the special case of a finite state space, as this is sufficient for the model in Chapter 5. For a more general treatment of the topic, the interested reader is referred to the literature, e.g., [GS06, Nor97, Res92, Ros96].

As the system is affected by random influences, the state of the system is a random object and its evolution can be modeled as a stochastic process. A stochastic or random process is a family $\mathcal{X} := \{X_t \mid t \in T\}$ of random variables indexed by some set $T$, i.e., there is an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and each $X_t : \Omega \to \mathbb{R}$ is $\mathcal{F}$-measurable. In our setting, the index $t \in T$ represents the time and $X_t$ models the state of the system at time $t$. Hence, $X_t(\omega)$ is an element of $S$ (for all $\omega \in \Omega$). As $S$ is finite, each $X_t$ is a discrete random variable. To match the formal definition of random variables, we assume that each state in $S$ is identified with an integer in $\mathbb{N}_0$. If the index set $T = \mathbb{N}_0$, $\mathcal{X}$ is called a discrete-time stochastic process and if $T = [0, \infty)$ (or $T = \mathbb{R}$), $\mathcal{X}$ is called a continuous-time stochastic process. In Sections 4.1.1 and 4.1.2, we review the theory of discrete-time and continuous-time homogeneous Markov chains, respectively, two important classes of stochastic processes. Our exposition is mainly based on [GS06, Nor97, Res92, Ros96].

---

1 [GS06, p. 213].
2 cf. [GS06, p. 214].
3 cf. [GS06, p. 214].
4 cf. [GS06, p. 213].
4.1 On homogeneous Markov chains

4.1.1 On discrete-time homogeneous Markov chains

Throughout this section, we assume that \( S = \{1, 2, \ldots, |S|\} \) with \(|S| < \infty\), i.e., we identify each state of the system with a unique integer number. Also, in accordance to the usual notation in the literature, we write \( X_n \) instead of \( X_t \) to denote the state of the system at time \( n \).

A discrete-time stochastic process \( X = \{X_n \mid n \in \mathbb{N}_0\} \) is a discrete-time Markov chain if it satisfies the Markov property:

\[
\mathbb{P}[X_n = i_n \mid X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}] = \mathbb{P}[X_n = i_n \mid X_{n-1} = i_{n-1}]
\]

for all \( n \in \mathbb{N} \), and all \( i_0, i_1, \ldots, i_n \in S \). For a discrete-time Markov chain, the subsequent state \( X_n \) thus only depends on the currently observed state \( X_{n-1} \). A discrete-time Markov chain \( X \) is called homogeneous if for all \( i, j \in S \), \( n \in \mathbb{N} \),

\[
\mathbb{P}[X_n = j \mid X_{n-1} = i] = \mathbb{P}[X_1 = j \mid X_0 = i],
\]

i.e., the “probabilistic behavior” of state transitions does not change over time. We use the abbreviation \( \text{DTHMC} \) for discrete-time homogeneous Markov chain. For the remainder of this section, let \( X = \{X_n \mid n \in \mathbb{N}_0\} \) denote a DTHMC.

The transition probabilities \( p(j \mid i) \) of \( X \) are defined by

\[
p(j \mid i) := \mathbb{P}[X_1 = j \mid X_0 = i] \quad \quad \quad \quad \quad \quad \quad \quad \quad i, j \in S,
\]

and the transition matrix \( P \in \mathbb{R}^{|S| \times |S|} \) is defined by

\[
P_{ij} := p(j \mid i) \quad \quad \quad \quad \quad \quad \quad \quad \quad i, j \in S.
\]

Evidently, \( P \) is a stochastic matrix. The “\( n \)-step transition probabilities” \( \mathbb{P}[X_{m+n} = j \mid X_m = i] \) satisfy the Chapman-Kolmogorov equations

\[
\mathbb{P}[X_{m+n+r} = j \mid X_m = i] = \sum_{k \in S} \mathbb{P}[X_{m+n} = k \mid X_m = i] \cdot \mathbb{P}[X_{m+n+r} = j \mid X_{m+n} = k],
\]

\((4.1)\)

\( ^{5} \) cf. [GS06, Definition 6.1.1].

\( ^{6} \) cf. [GS06, Definition 6.1.4].

\( ^{7} \) We adopt the notation \( p(j \mid i) \) from [Put94, p. 587].

\( ^{8} \) A matrix \( A \) is a stochastic matrix if each entry of \( A \) is non-negative and each row sum of \( A \) equals 1, cf. [GS06, Theorem 6.1.5].
for all \(i, j \in S\) and \(m, n, r \in \mathbb{N}_0\). These equations provide a formula for calculating \(\mathbb{P}[X_{m+n} = j \mid X_m = i]\) from the transition probabilities \(p(j \mid i)\), namely
\[
\mathbb{P}[X_{m+n} = j \mid X_m = i] = \sum_{k_1, \ldots, k_{n-1} \in S} p(k_1 \mid i) \cdot p(k_2 \mid k_1) \cdots p(j \mid k_{n-1}),
\]
(4.2)
for \(m, n \in \mathbb{N}_0, n \geq 2\). Thus, \(\mathbb{P}[X_{m+n} = j \mid X_m = i]\) is independent of \(m\) (for all \(n \in \mathbb{N}_0\)) and we write
\[
p^{(n)}(j \mid i) \quad \text{for} \quad \mathbb{P}[X_{m+n} = j \mid X_m = i] \quad n \in \mathbb{N}_0.
\]
The values \(p^{(n)}(j \mid i)\) are referred to as \emph{n-step transition probabilities}. It holds that
\[
p^{(n)}(j \mid i) = (P^n)_{ij}
\]
for \(i, j \in S, n \in \mathbb{N}_0\), i.e., the \(n\)-step transition probabilities are obtained from the \(n\)-th power of the transition matrix \(P\).

The \emph{initial distribution} \(p_0 \in \mathbb{R}^{|S|}\) of \(\mathcal{X}\) is defined by \((p_0)_i := \mathbb{P}[X_0 = i]\) for \(i \in S\) and we have
\[
\mathbb{P}[X_n = i] = (p_0^T P^n)_i
\]
(4.3)
for \(i \in S\) and \(n \in \mathbb{N}_0\), i.e., the transition matrix \(P\) and the initial distribution \(p_0\) determine the evolution of the DTHMC \(\mathcal{X}\).

\section*{Classification of states and chains}

A state \(i \in S\) is called \emph{recurrent} if
\[
\mathbb{P}[X_n = i \text{ for some } n \geq 1 \mid X_0 = i] = 1,
\]
(4.4)
and \(i\) is called \emph{transient} if the probability in (4.4) is strictly less than 1.\(^{13}\)

If the chain is started in a recurrent state \(i \in S\), \(i\) is visited infinitely

---

\(^9\) [GS06, Theorem 6.1.7].
\(^10\) [GS06, Theorem 6.1.7].
\(^11\) [GS06, Lemma 6.1.8].
\(^12\) cf. [GS06, p. 216].
\(^13\) cf. [GS06, Definition 6.2.1].
often (with probability 1), and – independently of the starting state – a
transient state is only visited a finite number of times (with probability
1).\textsuperscript{14} We have the following important result.\textsuperscript{15}

**Theorem 4.1.1** If the state space $S$ is finite, at least one state is recurrent.

Consequently (for finite $S$), the chain only visits recurrent states after a
finite number of steps (with probability 1).

For $i, j \in S$, $j$ is said to be accessible from $i$ (written $i \rightarrow j$) if there
exists an $n \in \mathbb{N}_0$ such that $p^{(n)}(j \mid i) > 0$, and $i$ and $j$ communicate
(written $i \leftrightarrow j$) if $i \rightarrow j$ and $j \rightarrow i$.\textsuperscript{16} As $i \rightarrow i$ (by definition) and
using (A.2) from page 172, we conclude that $\leftrightarrow$ is an equivalence relation
on $S$.\textsuperscript{17} Thus, $\leftrightarrow$ partitions $S$ into equivalence classes, which are called
**communicating classes.** As shown in the next theorem, all states in a
communicating class are either transient or recurrent.\textsuperscript{18}

**Theorem 4.1.2** Let $i \leftrightarrow j$. Then

(i) $i$ is transient $\iff$ $j$ is transient.

(ii) $i$ is recurrent $\iff$ $j$ is recurrent.

A communicating class $C \subseteq S$ is called recurrent (transient) if all states
in $C$ are recurrent (transient). A set of states $C \subseteq S$ is called closed if
$p(j \mid i) = 0$ for all $i \in C, j \notin C$.\textsuperscript{19} A state $i \in S$ is called absorbing if
the set $\{i\}$ is closed, i.e., $i$ is absorbing if and only if $p(i \mid i) = 1$.\textsuperscript{20} It
can be shown that each recurrent communicating class is closed and as a
consequence, there is a unique partition of the state space

$$S = T \cup C_1 \cup C_2 \cup \cdots \cup C_K$$

(4.5)

for some $K \in \mathbb{N}$, where $T$ is the set of transient states,\textsuperscript{21} and the $C_k$
$1 \leq k \leq K$, are closed recurrent communicating classes.\textsuperscript{22} If the chain

\textsuperscript{14} cf. proof of [Ros96, Proposition 4.2.3]; cf. also [Ber07, p. 177].

\textsuperscript{15} [GS06, Lemma 6.3.5].

\textsuperscript{16} cf. [Res92, p. 78, 79].

\textsuperscript{17} cf. also [Ros96, Proposition 4.2.1].

\textsuperscript{18} (i) is proven in [GS06, Theorem 6.3.2] and (ii) follows from (i) by contraposition.

\textsuperscript{19} cf. [GS06, Definition 6.3.3].

\textsuperscript{20} [Res92, p. 80].

\textsuperscript{21} $T$ can consist of several communicating classes, cf. [Res92, p. 99].

\textsuperscript{22} [GS06, Theorem 6.3.4].
is started in a state that belongs to some $C_k$, it never leaves $C_k$. In this case, $C_k$ can be considered to be the whole state space. On the other hand, if the chain is started in a state in $T$, it leaves $T$ after a finite number of steps and moves to some set $C_k$ where it remains ever after.\footnote{cf. [GS06, p. 225].}

**Example 4.1.3** Figure 4.1 illustrates an example of a DTHMC $\mathcal{X}$ with state space $\{1, 2, \ldots, 8\}$. The states are indicated by circles. The possible transitions are marked with arcs and the corresponding transition probabilities $p(j \mid i)$ are indicated next to the arcs. The communicating classes of $\mathcal{X}$ are

$$\{1\}, \{2\}, \{3, 4\}, \{5, 6, 7\}, \{8\}.$$  

The communicating classes $\{1\}, \{2\}, \{3, 4\}$ are transient and the communicating classes $\{5, 6, 7\}$ and $\{8\}$ are closed and recurrent. As a consequence, 8 is an absorbing state. \hfill ♦

![Figure 4.1 Example of a DTHMC.](image)

For a finite state space $S$, the closed recurrent communicating classes and the transient states of a DTHMC can be efficiently determined.\footnote{cf. the labeling algorithm in [Put94, Section A.3] which is based on [FL68].}

The DTHMC $\mathcal{X}$ is called *irreducible*, if the state space $S$ forms a single communicating class. For a finite state space $S$, the DTHMC $\mathcal{X}$ is called *unichain*\footnote{cf. [Put94, p. 589].} if $S$ consists of a single closed recurrent communicating class and a (possibly empty) set of transient states.
Stationary distribution, limiting matrix, long-term average cost

A vector $\pi \in \mathbb{R}^{|S|}$ is called stationary distribution\(^{26}\) of $X$ if $\pi$ is a probability distribution (i.e., $\pi_i \geq 0$ for all $i \in S$ and $\sum_{i \in S} \pi_i = 1$), and

$$\pi^T = \pi^T P.$$ \hfill (4.6)

If the initial distribution of $X$ (i.e., the distribution of $X_0$) is chosen to be a stationary distribution $\pi$, then we have (cf. (4.3)) $P[X_n = i] = (\pi^T P^n)_i = \pi_i$ for any $n \in \mathbb{N}_0$.\(^{27}\) If $\pi \in \mathbb{R}^{|S|}$ satisfies (4.6) and $i \in S$ is transient, then $\pi_i = 0$.\(^{28,29}\) As a consequence, the condition (4.6) on a stationary distribution decomposes into an independent equality system for each closed recurrent communicating class $C_k$ (cf. (4.5)).\(^{30}\) The following theorem gives a result on the existence of stationary distributions.\(^{31}\)

**Theorem 4.1.4** Assume $S$ is finite and let $X$ denote a unichain DTHMC with closed recurrent communicating class $C$ and transient states $T$. Then $X$ has a unique stationary distribution $\pi$ and $\pi$ satisfies $\pi_i > 0$ for $i \in C$ and $\pi_i = 0$ for $i \in T$.

\(^{26}\) cf. [GS06, Definition 6.4.1].

\(^{27}\) cf. [GS06, p. 227].

\(^{28}\) For a transient state $i \in S$, we have $p^{(n)}(i \mid j) \to 0$ as $n \to \infty$ for all $j \in S$ ([GS06, Corollary 6.2.5]). As $\pi^T P^n = \pi^T$ for all $n$, i.e., $\pi_i = \sum_{j \in S} \pi_j (P^n)_{ji} = \sum_{j \in S} \pi_j p^{(n)}(i \mid j)$ for all $n$, the claim follows by contradiction.

\(^{29}\) The converse is not true in general. Consider a DTHMC with $S = \{1, 2\}$ and transition probabilities $p(1 \mid 1) = 1$ and $p(2 \mid 2) = 1$, i.e., states 1 and 2 are absorbing and thus recurrent. However, $\pi = (1, 0)^T$ is a stationary distribution with $\pi_2 = 0$.

\(^{30}\) Decompose $S$ as in (4.5) and let $i$ be a recurrent state belonging to $C_{k(i)}$. Then

$$\pi_i = \sum_{j \in T} \pi_j p(i \mid j) + \sum_{j \in C_{k(i)}} \pi_j p(i \mid j) + \sum_{k \neq k(i)} \sum_{j \in C_k} \pi_j p(i \mid j) = 0.$$ \hfill (4.6)

Further, for $i \in T$, we have

$$0 = \pi_i = \sum_{j \in T} \pi_j p(i \mid j) + \sum_{k} \sum_{j \in C_k} \pi_j p(i \mid j) = 0.$$

\(^{31}\) cf. [GS06, Theorem 6.4.3 and Lemma 6.3.5] in combination with the above arguments that a stationary distribution is zero on transient states and that the condition (4.6) can be decomposed. Cf. also [Put94, Proposition 8.8.5].
The limiting matrix\(^{32}\) \(P^*\) (of \(P\)) is defined element-wise by

\[
P^*_{ij} := \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} p^{(n)}(j \mid i).
\]

(4.7)

It can be shown that the limit in (4.7) exists for all \(i, j \in S\), i.e., \(P^*\) is well-defined, and that the limiting matrix \(P^*\) has the following properties.\(^{33}\)

**Theorem 4.1.5** Let \(S\) be finite. Then

(i) \(P^*\) is a stochastic matrix.

(ii) \(P^*\) satisfies

\[
P^* = PP^* = P^*P = P^*P^*.
\]

As \(P^*\) is stochastic, each row of \(P^*\) satisfies (4.6) due to (ii) of the previous theorem, i.e., each row of \(P^*\) is a stationary distribution. Thus, the following theorem results from Theorems 4.1.4 and 4.1.5.\(^{34}\)

**Theorem 4.1.6** Assume \(S\) is finite and let \(\mathcal{X}\) denote a unichain DTHMC. Then, the rows of \(P^*\) are identical and correspond to the unique stationary distribution of \(\mathcal{X}\).

We conclude this section with the following theorem about the long-term average cost of a DTHMC, where a cost is associated with each state.\(^{35,36}\)

\(^{32}\) cf. [Put94, p. 592].

\(^{33}\) [Doo53, Theorem 2.1 in Chapter V].

\(^{34}\) cf. also [Put94, p. 593, 594].

\(^{35}\) [Res92, Corollary 2.12.5] gives the same result with slightly different assumptions.

\(^{36}\) Let \(i_r \in S\) be a recurrent state. Starting from any \(i \in S\), \(i_r\) is visited after a finite number of steps with probability 1. Thus, the proof of the ergodic theorem [Nor97, Theorem 1.10.2] yields

\[
P[ lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} g(X_n) = \sum_{j \in S} g(j)\pi_j \mid X_0 = i] = 1 \quad \forall i \in S.
\]

As \(S\) is finite, we have \(\frac{1}{N} \sum_{n=0}^{N-1} g(X_n(\omega)) \leq \max_{j \in S} |g(j)| < \infty\) for all \(\omega \in \Omega\) and the theorem follows from the dominated convergence theorem, [Dur10, Theorem 1.6.7].
Theorem 4.1.7 Let $S$ be finite and let $g : S \to \mathbb{R}$ be a function on $S$. Assume the DTHMC $X$ is unichain and let $\pi$ denote its (unique) stationary distribution. Then

$$\lim_{N \to \infty} \frac{1}{N} \mathbb{E} \left[ \sum_{n=0}^{N-1} g(X_n) \mid X_0 = i \right] = \sum_{j \in S} g(j) \pi_j \quad i \in S.$$ 

In Appendix A.1, we provide some additional properties of discrete-time homogeneous Markov chains.

4.1.2 On continuous-time homogeneous Markov chains

Throughout this section, we again assume that $S = \{1, 2, \ldots, |S|\}$ with $|S| < \infty$. To avoid subtleties that arise in the analysis of continuous-time stochastic processes $X = \{X_t \mid t \in [0, \infty)\}$, all processes are assumed to be right-continuous, i.e., for all $\omega \in \Omega$ and $t \geq 0$, there exists an $\epsilon > 0$ such that $X_s(\omega) = X_t(\omega)$ for all $t \leq s \leq t + \epsilon$.\textsuperscript{37} For a right-continuous stochastic process $X = \{X_t \mid t \in [0, \infty)\}$, the probability of any event depending on $X$ is uniquely determined by its finite-dimensional distributions

$$\mathbb{P}[X_{t_0} = i_0, X_{t_1} = i_1, \ldots, X_{t_n} = i_n] \quad (4.8)$$

for $n \in \mathbb{N}_0$, $0 \leq t_0 \leq t_1 \leq \cdots \leq t_n$ and $i_0, \ldots, i_n \in S$.\textsuperscript{38} The jump times\textsuperscript{39} $J_n$, $n \in \mathbb{N}_0$, of $X$ are defined by

$$J_0 := 0$$

$$J_{n+1} := \inf \{t \geq J_n \mid X_t \neq X_{J_n} \} \quad n \in \mathbb{N}_0$$

with the convention that $\inf \emptyset = \infty$ and the holding times\textsuperscript{40} $H_n$, $n \in \mathbb{N}$, are defined by

$$H_n := \begin{cases} J_n - J_{n-1} & \text{if } J_{n-1} < \infty \\ \infty & \text{otherwise} \end{cases} \quad n \in \mathbb{N}.$$
We restrict our attention to a special type of continuous-time stochastic processes, namely to continuous-time homogeneous Markov chains. Following [Nor97], we say that a random variable \( Z : \Omega \to [0, \infty] \) is exponentially distributed with parameter \( \lambda \), \( 0 \leq \lambda < \infty \), if for all \( t \geq 0 \)

\[
P[Z > t] = e^{-\lambda t} \quad (4.9)
\]

and we write \( Z \sim \text{EXP}(\lambda) \). If \( \lambda = 0 \), \( P[Z = \infty] = 1 \); if \( \lambda > 0 \), \( Z \) has density function

\[
f_Z(t) = \lambda e^{-\lambda t} \mathbf{1}_{\{t \geq 0\}}.
\]

The expected value of \( Z \sim \text{EXP}(\lambda) \) is

\[
E[Z] = \frac{1}{\lambda} \quad (with \text{the convention that } \frac{1}{0} = \infty).
\]

The definition of a continuous-time homogeneous Markov chain is given next.

**Definition 4.1.8** Let \( S \) be finite. A continuous-time stochastic process \( \mathcal{X} = \{X_t \mid t \in [0, \infty)\} \) is called continuous-time homogeneous Markov chain (CTHMC) if

(i) the holding times of \( \mathcal{X} \) are independent random variables,

(ii) the holding times in state \( i \in S \) are (independent and) identically distributed, \( \sim \text{EXP}(\alpha_i) \), for some \( 0 \leq \alpha_i < \infty \),

(iii) the jump chain, i.e., the discrete-time stochastic process \( \mathcal{X}^\text{jump} := \{X^{\text{jump}}_n \mid n \in \mathbb{N}_0\} \) with \( X^{\text{jump}}_n := X_{J_n} \), is a discrete-time homogeneous Markov chain with transition matrix \( \mathbf{P}^{\text{jump}} = (p(j \mid i))_{i,j \in S} \) that satisfies \( p(i \mid i) = 0 \) if \( \alpha_i \neq 0 \) and \( p(i \mid i) = 1 \) if \( \alpha_i = 0 \).

\( \alpha_i \) is the rate at which the chain leaves state \( i \in S \) and we thus refer to \( \alpha_i \) as the rate of leaving \( i \).\(^{44}\) When leaving \( i \), the chain moves to state \( j \in S \) with probability \( p(j \mid i) \), i.e., \( \alpha_i p(j \mid i) \) is the rate of moving from \( i \) to \( j \). We refer to \( \alpha_i p(j \mid i) \) as the transition rate from \( i \) to \( j \).\(^{45}\) We further call \( p(j \mid i) \) the jump probability from \( i \) to \( j \).\(^{46}\) The rates of leaving \( \alpha_i, i \in S \), and the transition matrix \( \mathbf{P}^{\text{jump}} \) of the jump chain are the characteristic data for a CTHMC.

For the remainder of this section, let \( \mathcal{X} = \{X_t \mid t \in [0, \infty)\} \) denote a CTHMC. An extract of a possible realization of the state of \( \mathcal{X} \) and the state of its jump chain \( \mathcal{X}^{\text{jump}} \) is shown in Figure 4.2.

\(^{41}\) cf. [Nor97, p. 70].

\(^{42}\) cf. [Nor97, p. 70].

\(^{43}\) The definition is based on [Nor97, Theorem 2.8.2].

\(^{44}\) cf. [Nor97, p. 89].

\(^{45}\) cf. [Ros96, p. 233].

\(^{46}\) According to the notion of the jump matrix in [Nor97, p. 87].
4.1 On homogeneous Markov chains

A state \( i \in S \) is called absorbing (for \( \mathcal{X} \)) if \( \alpha_i = 0 \) and stable if \( 0 < \alpha_i < \infty \). If the chain reaches an absorbing state, it remains in this state forever, because the corresponding holding time is infinite. Observe that \( i \in S \) is absorbing (for \( \mathcal{X} \)) if and only if \( i \) is absorbing for the jump chain \( \mathcal{X}^{\text{jump}} \).

The CTHMC \( \mathcal{X} \) possesses the Markov property

\[
P[X_{t_{n+1}} = i_{n+1} \mid X_{t_0} = i_0, X_{t_1} = i_1, \ldots, X_{t_n} = i_n] = P[X_{t_{n+1}} = i_{n+1} \mid X_{t_n} = i_n]
\]

for all \( n \in \mathbb{N}_0, 0 \leq t_0 < t_1 < \cdots < t_{n+1} \), and \( i_0, i_1, \ldots, i_{n+1} \in S \). Thus, the state \( X_{t_{n+1}} \) at time \( t_{n+1} \) only depends on the currently observed state \( X_{t_n} \). In addition, \( \mathcal{X} \) is homogeneous, i.e.,

\[
P[X_t = j \mid X_s = i] = P[X_{t-s} = j \mid X_0 = i]
\]

for all \( i, j \in S \) and all \( 0 \leq s \leq t \). Hence, the “probabilistic evolution” of the chain, given it is in some state \( i \) at time \( s \), does not depend on the time of observation \( s \).

Further, the generator (matrix) \( Q \in \mathbb{R}^{|S|\times|S|} \) of the CTHMC \( \mathcal{X} \) is

\[\text{Figure 4.2} \quad \text{Extract of a possible evolution of the state of a CTHMC} \: \mathcal{X} \: \text{and the state of its jump chain} \: \mathcal{X}^{\text{jump}}.\]
defined as

\[ Q_{ii} := -\alpha_i \]
\[ Q_{ij} := \alpha_i p(j \mid i) \quad j \neq i, \quad (4.10) \]
i.e., the diagonal entries of \( Q \) are the negative values of the rates of leaving and the off-diagonal entries are the transition rates. Note that from the knowledge of the generator \( Q \), the rates of leaving \( \alpha_i \) as well as the jump probabilities \( p(j \mid i) \) can be obtained. The “probabilistic behavior” of \( \mathcal{X} \) is governed by its generator \( Q \). More precisely, the family of matrices \( P(t), t \geq 0 \), defined by

\[(P(t))_{ij} := \mathbb{P}[X_t = j \mid X_0 = i] \]

for \( i, j \in S \) is given by

\[ P(t) = e^{tQ} \]

for \( t \geq 0 \),\(^{51}\) where \( e^{tQ} := \sum_{k=0}^{\infty} \frac{1}{k!}(tQ)^k, \quad t \geq 0 \).\(^{52}\) Observe that the finite-dimensional distributions \((4.8)\) of \( \mathcal{X} \) can be obtained from \( P(t) \).

**Stationary distribution, long-term average cost**

A vector \( \pi \in \mathbb{R}^{|S|} \) is called *stationary distribution*\(^{53}\) of the CTHMC \( \mathcal{X} = \{X_t \mid t \in [0, \infty)\} \) if \( \pi \) is a probability distribution (i.e., \( \pi_i \geq 0 \) for all \( i \in S \) and \( \sum_{i \in S} \pi_i = 1 \)), and

\[ \pi^T = \pi^T P(t) \quad \text{for all } t \geq 0. \]

As in the discrete-time case, if the chain is started according to a stationary distribution \( \pi \), i.e., if \( \mathbb{P}[X_0 = i] = \pi_i \) for all \( i \in S \), then \( \mathbb{P}[X_t = i] = \pi_i \) for any \( t \geq 0 \) and \( i \in S \).\(^{54}\)

---

\(^{51}\) [Nor97, Theorem 2.8.2 and Theorem 2.1.1].

\(^{52}\) \( e^{tQ}, \quad t \geq 0 \), is a well-defined stochastic matrix, cf. [Nor97, Section 2.10 and Theorem 2.1.2].

\(^{53}\) cf. [Res92, p. 392].

\(^{54}\) [Res92, Proposition 5.5.1].
For finite $S$ and a vector $\pi \in \mathbb{R}^{|S|}$, it can be shown that
\[
\pi^T = \pi^T P(t), \forall t \geq 0 \text{ if and only if } \pi^T Q = 0.
\]

For $\pi \in \mathbb{R}^{|S|}_{\geq 0}$ with $\pi^T Q = 0$, we have $\pi_i = 0$ for all states $i \in S$ that are transient for the jump chain $\mathcal{X}^{\text{jump}}$ of $\mathcal{X}$. Consequently, the condition $\pi^T Q = 0$ on a stationary distribution of $\mathcal{X}$ decomposes into an independent equality system for each closed recurrent communicating class of the jump chain $\mathcal{X}^{\text{jump}}$. The following theorem gives a result on the existence of stationary distributions.

**Theorem 4.1.9** Let $S$ be finite. Let $\mathcal{X}$ denote a CTHMC and assume that its jump chain $\mathcal{X}^{\text{jump}}$ is unichain with closed recurrent communicating class $C$ and transient states $T$. Then $\mathcal{X}$ has a unique stationary distribution $\pi$ and $\pi$ satisfies $\pi_i > 0$ for $i \in C$ and $\pi_i = 0$ for $i \in T$.

---

55 cf. the beginning of the proof of [Nor97, Theorem 3.5.5]: as $\frac{d}{dt} P(t) = QP(t)$ for all $t \geq 0$ ([Nor97, Theorem 2.1.1]), it follows that (using $S$ is finite)
\[
\frac{d}{dt} \pi^T P(t) = \pi^T (\frac{d}{dt} P(t)) = \pi^T QP(t) .
\]

Thus, if $\pi^T Q = 0$, $\pi^T P(t)$ is constant, i.e., $\pi^T P(t) = \pi^T P(0) = \pi^T$ for all $t \geq 0$. On the other hand, if $\pi^T P(t) = \pi^T$ for all $t \geq 0$, we have from (4.11) $\pi^T QP(t) = \frac{d}{dt} \pi^T P(t) = 0$ for all $t \geq 0$ and for $t = 0$ we obtain $\pi^T Q = 0$.

56 Define $\nu \in \mathbb{R}^{|S|}$ by $\nu_i := \alpha_i \pi_i$. By [Nor97, Theorem 3.5.1], we have $\nu^T = \nu^T P^{\text{jump}}$. For a state $i$ that is transient for $\mathcal{X}^{\text{jump}}$, we have $\alpha_i > 0$ ([Nor97, Theorem 3.4.1 and Theorem 3.4.2]). Now suppose $\pi_i > 0$. Then $\nu_i > 0$, which is a contradiction to the fact that a stationary distribution of a DTHMC is zero on transient states.

57 By a similar argument as in Footnote 30 on page 67.

58 cf. [Res92, Theorem 5.5.3] together with the above arguments that a stationary distribution is zero on transient states and that the condition $\pi^T Q = 0$ can be decomposed.
The following theorem provides a result on the long-term average cost of a CTHMC, where a cost is accumulated when being in a state.\footnote{\cite[p. 397]{Res92} gives the same result with slightly different assumptions.} \footnote{\cite[p. 438]{Res92}.}

**Theorem 4.1.10** Let $S$ be finite and let $g : S \rightarrow \mathbb{R}$ be a function on $S$. Let $X$ denote a CTHMC whose jump chain is unichain and let $\pi$ denote the (unique) stationary distribution of $X$. Then

$$
\lim_{T \to \infty} \frac{1}{T} \mathbb{E} \left[ \int_0^T g(X_t) dt \mid X_0 = i \right] = \sum_{j \in S} g(j) \pi_j \quad i \in S.
$$

**Uniformization**

Let $X = \{X_t \mid t \in [0, \infty]\}$ denote a CTHMC. In the following, an auxiliary discrete-time homogeneous Markov chain is defined that proves to be useful for solving continuous-time Markov decision processes.

As $S$ is finite, there exists $\eta \in \mathbb{R}_{>0}$, such that $\alpha_i \leq \eta$ for all $i \in S$. The auxiliary DTHMC $\tilde{X}^\eta := \{\tilde{X}_n^\eta \mid n \in \mathbb{N}_0\}$ is defined via its transition probabilities

$$
\tilde{p}(j \mid i) := \begin{cases} 
\frac{\alpha_i}{\eta} p(j \mid i) & \text{if } j \neq i \\
1 - \frac{\alpha_i}{\eta} & \text{if } j = i
\end{cases}
$$

(4.12)

for $i, j \in S$.\footnote{\cite[p. 398]{Res92}.} When being in state $i \in S$, the DTHMC $\tilde{X}^\eta$ leaves $i$ with probability $\frac{\alpha_i}{\eta}$ and remains in $i$ with probability $1 - \frac{\alpha_i}{\eta}$. If $\alpha_i \notin \{0, \eta\}$, the number of transitions until state $i$ is left is thus given by a geometric random variable with parameter $\frac{\alpha_i}{\eta}$ and its expected value is $\frac{\eta}{\alpha_i}$.\footnote{\cite[p. 121]{Ste09}.} If transitions of the discrete-time chain are assumed to occur every $\frac{1}{\eta}$ time

As $S$ is finite, we have $\frac{1}{T} \int_0^T g(X_t(\omega)) dt \leq \max_{j \in S} |g(j)| < \infty$ for all $\omega \in \Omega$ and the theorem follows from the dominated convergence theorem \cite[Theorem 1.6.7]{Dur10}.
units, the chain remains in \( i \) for \( \frac{1}{\alpha_i} \) time units in expectation, which equals the expected holding time of \( \mathcal{X} \) in \( i \). When leaving state \( i \), the process moves to state \( j \) with probability \( \frac{\alpha_i}{\eta} \cdot p(j \mid i) \), which is the probability of leaving \( i \) times the probability that \( \mathcal{X} \) jumps to \( j \).

It is now shown how the original continuous-time chain \( \mathcal{X} \) can be “reconstructed” from the auxiliary discrete-time chain \( \tilde{\mathcal{X}}^\eta \). More precisely, a new continuous-time stochastic process \( \mathcal{X}^\eta \) is constructed that is “equivalent” to \( \mathcal{X} \). For the construction, a special type of CTHMC is needed, namely a Poisson process. Following [Nor97], we say that a continuous-time stochastic process \( N := \{N_t \mid t \in [0, \infty)\} \) is a (homogeneous) Poisson process of rate \( \alpha \), \( 0 < \alpha < \infty \), if its holding times are independent exponential random variables with parameter \( \alpha \) and its jump chain \( \{N_n^{\text{jump}} \mid n \in \mathbb{N}_0\} \) is given by \( N_n^{\text{jump}} = n \) for all \( n \in \mathbb{N}_0 \).

Now let \( \mathcal{N} := \{N_t \mid t \in [0, \infty)\} \) denote a Poisson process of rate \( \eta \). The new continuous-time stochastic process \( \mathcal{X}^\eta := \{X_t^\eta \mid t \in [0, \infty)\} \) is defined by

\[
X_t^\eta := \tilde{X}^\eta_{N(t)}
\]

for all \( t \geq 0 \), i.e., the “jump times” of \( \mathcal{X}^\eta \) are specified by the Poisson process \( \mathcal{N} \) and the state of \( \mathcal{X}^\eta \) changes according to the transitions of the discrete-time chain \( \tilde{\mathcal{X}}^\eta \).\(^{63}\) Note that fictitious transitions are allowed from a state back to itself,\(^{64}\) because \( \tilde{p}(i \mid i) > 0 \) whenever \( \alpha_i \neq \eta \). Allowing these fictitious transitions, the process \( \mathcal{X}^\eta \) can be interpreted as having uniform “rates of leaving” (each of size \( \eta \)) and hence \( \mathcal{X}^\eta \) is referred to as uniformization of \( \mathcal{X} \).\(^{65}\) It can be shown that

\[
P[X_t^\eta = j \mid X_0^\eta = i] = P[X_t = j \mid X_0 = i]
\]

for all \( t \geq 0 \) and \( i,j \in S \),\(^{66}\) i.e., the “probabilistic evolution” of \( \mathcal{X}^\eta \) and \( \mathcal{X} \) is identical.

For distinct states \( i, j \in S, i \neq j \), we have \( \tilde{p}(j \mid i) = 0 \Leftrightarrow p(j \mid i) = 0 \). From Theorem A.1.2 it follows (using \( |S| < \infty \)) that the auxiliary

\(^{63}\) cf. [Res92, p. 436].
\(^{64}\) cf. [Put94, p. 562].
\(^{65}\) cf. [Put94, p. 562].
\(^{66}\) [Res92, Example 5.1.3 (and Section 5.10)] and [Ste09, p. 361].
\(^{67}\) If \( \tilde{p}(j \mid i) = 0 \) then either \( p(j \mid i) = 0 \) or \( \alpha_i = 0 \). But \( \alpha_i = 0 \) implies \( p(j \mid i) = 0 \) because \( \rho(i \mid i) = 1 \) due to Definition 4.1.8(iii). The other direction follows from the definition of \( \tilde{p}(j \mid i) \).
DTHMC $\tilde{X}^\eta$ and the jump chain $X^{\text{jump}}$ of $X$ have the same recurrent and transient states and as a consequence, we have:

\begin{equation}
\text{Theorem 4.1.11} \hspace{1cm} \text{Assume } |S| < \infty. \text{ Then, the auxiliary DTHMC } \tilde{X}^\eta \hspace{1cm} \text{(defined via (4.12)) is unichain if and only if the jump chain } X^{\text{jump}} \text{ of } X \hspace{1cm} \text{is unichain.}
\end{equation}

Observe that the stationary distributions of the CTHMC $X$ and the DTHMC $\tilde{X}^\eta$ are identical. Assuming that the jump chain of $X$ is unichain, we conclude from Theorems 4.1.10, 4.1.11, and 4.1.7 (for any function $g : S \to \mathbb{R}$ and for all $i \in S$) that

\[ \lim_{T \to \infty} \frac{1}{T} \mathbb{E}\left[ \int_0^T g(X_t) \, dt \mid X_0 = i \right] = \lim_{N \to \infty} \frac{1}{N} \mathbb{E}\left[ \sum_{n=0}^{N-1} g(\tilde{X}^\eta_n) \mid \tilde{X}^\eta_0 = i \right]. \]

Interpreting $g(j)$ as the cost rate at which the cost are accumulated when being in state $j \in S$ (and assuming that transitions of $\tilde{X}^\eta$ occur after 1 time unit), the long-term average cost of the CTHMC $X$ are equal to the long-term average cost of the DTHMC $\tilde{X}^\eta$.

## 4.2 On Markov decision processes

In this section, we summarize the theory of Markov decision processes, a modeling framework for sequential decision making in a stochastic environment. We use the abbreviation $MDP$ for Markov decision process.

The framework of Markov decision processes is introduced in Section 4.2.1. It is shown how a continuous-time MDP can be transformed into an (equivalent) discrete-time MDP. The equivalence, which is important for solving continuous-time MDPs, is discussed in Section 4.2.4. In Sections 4.2.2 and 4.2.3, the theory and algorithms of (unichain) discrete-time MDPs is summarized. Our exposition in this section is mainly based on [Put94] and partly on [Ber07] and most of the notions and definitions closely follow the lines of [Put94].

---

\footnote{cf. also [Ber07, p. 315, 316] in the context of semi-Markov decision processes.}

\footnote{The transition matrix $\tilde{P}$ of $\tilde{X}^\eta$ is given by $\tilde{P} = I + \frac{1}{\eta}Q$. Hence for any vector $\pi \in \mathbb{R}^{|S|}$ we have $\pi^T(\tilde{P} - I) = \frac{1}{\eta} \pi^T Q$ and thus $\pi^T \tilde{P} = \pi^T \Leftrightarrow \pi^T Q = 0$.}
4.2 On Markov decision processes

4.2.1 Discrete-time and continuous-time MDPs

Again, we consider a system whose state changes over time, due to some random influences. We assume that the state of the system is continuously observed. At (random) points in time, called decision epochs, the system can be influenced by the control

\(^{70}\) (or action) of a decision maker. The chosen control is applied until the next decision epoch. The goal of the decision maker is to optimally control the system.

As before, the set of possible states is denoted by \(S\) and is referred to as state space. For an observed state \(x \in S\), the set of (allowable) controls is denoted by \(U(x)\) and the set of (all) controls is \(U := \bigcup_{x \in S} U(x)\). Throughout Section 4.2, we assume that the set of states \(S\) and the set of controls \(U\) are finite as this is sufficient for the application in Chapter 5.

Remark 4.2.1 We allow \(S\) to be a “descriptive characterization” of the possible states of the system, i.e., \(S\) is a priori not a subset of \(\mathbb{N}_0\) (cf. the state description of the application in Chapter 5). However, we fix an arbitrary enumeration of \(S\) and also write \(x\) for the index of state \(x\) in this enumeration. Then, any real-valued function \(h\) on \(S\), i.e., \(h : S \to \mathbb{R}, x \mapsto h(x)\), can be identified with a vector in \(\mathbb{R}^{\mid S\mid}\) whose \(x\)-th component (according to the above enumeration of \(S\)) equals \(h(x)\). In the following, we thus refer to real-valued functions on \(S\) as vectors in \(\mathbb{R}^{\mid S\mid}\). The enumeration of \(S\) is also used to define transition matrices (cf. Section 4.2.2).

A decision rule \(\mu\) specifies how to choose a control at a given decision epoch. We mainly focus on deterministic Markovian decision rules where controls are chosen deterministically and only depend on the currently observed state of the system. A deterministic Markovian decision rule is a function \(\mu : S \to U, x \mapsto \mu(x) \in U(x)\) for all \(x \in S\). Let \(D^{\text{MD}}\) denote the set of all deterministic Markovian decision rules.

More generally, controls may be chosen randomly and/or history dependent. A randomized decision rule defines a probability distribution on the set \(U(x)\) of controls for each state \(x \in S\). A history dependent decision rule can depend on the previously visited states, the sojourn times in these states, and the chosen controls.

\(^{70}\) We use the term “control”, according to the literature in control theory, cf. e.g. [Ber07].
Let $\tau_n$, $n \in \mathbb{N}_0$, denote the (random) decision epochs and we assume that $\tau_0 := 0$. A policy $\pi$ specifies a sequence of decision rules, i.e., $\pi = \{\mu_0, \mu_1, \ldots\}$ where $\mu_n$, $n \in \mathbb{N}_0$, is the decision rule that is applied at decision epoch $\tau_n$. The set of all policies is denoted by $\Pi$ and the set of policies with deterministic Markovian decision rules is denoted by $\Pi^{MD}$. A policy $\pi = \{\mu, \mu, \ldots\} \in \Pi$ with the same decision rule at each decision epoch is called stationary and is denoted by $\mu$. Observe that the decision rule $\mu$ of a stationary policy is Markovian.

Under appropriate assumptions – which are satisfied in the application in Chapter 5 – it suffices to consider deterministic Markovian decision rules (and deterministic stationary policies) in the setting of average expected cost over an infinite time horizon (cf. Theorems 4.2.10 and 4.2.17).

It is assumed that the transitions to a new state occur at (a subset of) the decision epochs. As the system is affected by random influences, the evolution of its state can be modeled by a stochastic process. Given a policy $\pi \in \Pi$, the continuous-time stochastic process\textsuperscript{71} that models the evolution of the state in continuous time is referred to as natural process and is denoted by $X^\text{nat}_\pi := \{X^\text{nat}_{\pi,t} \mid t \geq 0\}$. The process that observes the state of the system at the decision epochs $\tau_n$ only is referred to as embedded decision process and is denoted by $X_\pi := \{X_{\pi,n} \mid n \in \mathbb{N}_0\}$, i.e., for $n \in \mathbb{N}_0$ we have

$$X_{\pi,n} = X^\text{nat}_{\pi,\tau_n}.$$ 

The evolution of the state of Markov decision processes (MDPs) is defined by jump probabilities, which determine the behavior of $X_\pi$, and the times between decision epochs. Given that the embedded decision process is in state $x$ at the current decision epoch and control $u$ is chosen, it is in state $\bar{x}$ at the next decision epoch with (conditional) jump probability $p(\bar{x} \mid x, u)$. The evolution of $X_\pi$ thus only depends on the observed state $x$ and the chosen control $u$ at the current decision epoch. Consequently, for MDPs, $X_\pi$ is referred to as embedded Markov decision process. For a stationary policy $\mu = \{\mu, \mu, \ldots\} \in \Pi$, the embedded Markov decision process $X_\mu := \{X_{\mu,n} \mid n \in \mathbb{N}_0\}$ is a discrete-time homogeneous Markov chain (cf. Section 4.1.1). $X_\mu$ is referred to as embedded Markov chain corresponding to $\mu$. If the stationary policy $\mu$ is deterministic, the transition

\textsuperscript{71} which we assume to be right-continuous (cf. Section 4.1.2).
probabilities of $X_\mu$ are given by $p(\bar{x} \mid x, \mu(x))$ for all $x, \bar{x} \in S$.

The time until the next decision epoch is governed by the (conditional) probability distribution $F(\cdot \mid x, u)$ that depends on the state $x$ at the current decision epoch and the chosen control $u$. We consider two particular forms of $F(\cdot \mid x, u)$.

- If the times between decision epochs are deterministic and identical for all states and controls, i.e., if for some fixed $\Delta > 0$,

$$F(t \mid x, u) = \begin{cases} 0 & 0 \leq t < \Delta \\ 1 & t \geq \Delta \end{cases} \quad x \in S, u \in U(x),$$

the corresponding process is called *discrete-time Markov decision process*. For a discrete-time MDP, we also refer to the jump probabilities $p(\bar{x} \mid x, u)$ as *transition probabilities* (in analogy to the notions of discrete-time homogeneous Markov chains).

An extract of a possible realization of the natural process $X^{\text{nat}}_\pi$ and the embedded Markov decision process $X_\pi$ of a discrete-time MDP, given a policy $\pi = \{\mu_0, \mu_1, \ldots\} \in \Pi^\text{MD}$, is shown in Figure 4.3.

- If the times between decision epochs are exponentially distributed with parameters $0 \leq \alpha_{x,u} < \infty$, $x \in S, u \in U(x)$, i.e., if

$$F(t \mid x, u) = 1 - e^{-\alpha_{x,u} t} \quad x \in S, u \in U(x), t \geq 0,$

the corresponding process is called *continuous-time Markov decision process*.

An extract of a possible realization of the natural process $X^{\text{nat}}_\pi$ and the embedded Markov decision process $X_\pi$ of a continuous-time MDP, given a policy $\pi = \{\mu_0, \mu_1, \ldots\} \in \Pi^\text{MD}$, is shown in Figure 4.4.

For a continuous-time MDP, we assume that the state changes at each decision epoch, i.e., we assume that $p(x \mid x, u) = 0$ if $\alpha_{x,u} \neq 0$. We also require that $p(x \mid x, u) = 1$ when $\alpha_{x,u} = 0$. In the latter case, upon reaching state $x$, the process remains in $x$ forever (when control $u$ is applied) and there is no subsequent decision epoch.

72 If the stationary policy $\mu$ is randomized and $q_{\mu(x)} \in \mathbb{R}^{|U(x)|}$ denotes the corresponding probability distribution on $U(x)$ for $x \in S$ (i.e., $u \in U(x)$ is chosen with probability $q_{\mu(x)}(u)$), the transition probabilities of $X_\mu$ are given by

$$\sum_{u \in U(x)} q_{\mu(x)}(u)p(\bar{x} \mid x, u)$$

for all $x, \bar{x} \in S$, cf. [Put94, p. 22].

73 cf. [Put94, p. 533].
Thus, the natural process $X_{\mu}^{\text{nat}}$ corresponding to a deterministic stationary policy $\mu \in \Pi^{\text{MD}}$ is a continuous-time homogeneous Markov chain (CTHMC; cf. Section 4.1.2) with rates of leaving $\alpha_{x,\mu}(x)$ and whose jump chain is $X_{\mu}$, the embedded Markov chain corresponding to $\mu$.

In analogy to the notions of CTHMCs, we refer to $\alpha_{x,u}$ as the rate of leaving state $x$ when choosing control $u$. The transition rate from $x$ to $\bar{x}$ when choosing control $u$ is defined by

$$\alpha(\bar{x} \mid x, u) := \alpha_{x,u} \cdot p(\bar{x} \mid x, u). \quad (4.14)$$

Further, we refer to the time until the next decision epoch, given the system is in state $x \in S$ and control $u \in U(x)$ is chosen at the current decision epoch, as holding time $H_{x,u}$ in state $x$ when control $u$ is chosen.

As $\sum_{\bar{x} \in S} \alpha(\bar{x} \mid x, u) = \alpha_{x,u}$ by (4.14), we have for all $x, \bar{x} \in S$ and $u \in U(x)$ with $\alpha_{x,u} \neq 0$

$$p(\bar{x} \mid x, u) = \frac{\alpha(\bar{x} \mid x, u)}{\sum_{x' \in S} \alpha(x' \mid x, u)}, \quad (4.15)$$

For a deterministic stationary policy $\mu$ and $\bar{x} \neq x$, $\alpha(\bar{x} \mid x, \mu(x))$ corresponds to the $(x, \bar{x})$-entry of the generator matrix of $X_{\mu}^{\text{nat}}$, cf. (4.10).
4.2 On Markov decision processes

Figure 4.4: Extract of a possible evolution of the natural process and the embedded Markov decision process of a continuous-time MDP (for a fixed policy \( \pi = \{\mu_0, \mu_1, \ldots \} \in \Pi^{\text{MD}} \)).

i.e., the behavior of a continuous-time MDP is fully characterized by the knowledge of either the transition rates \( \alpha(x, u) \) or the jump probabilities \( p(x \mid x, u) \) and the rates of leaving \( \alpha_{x,u} \).

A discrete-time MDP is denoted by \( \mathcal{M}^{\text{discr}} \) and is characterized by

\[
\mathcal{M}^{\text{discr}} = \{S, U(x), p(\cdot \mid x, u), g(x, u)\},
\]

and a continuous-time MDP is denoted by \( \mathcal{M}^{\text{cont}} \) and is characterized by

\[
\mathcal{M}^{\text{cont}} = \{S, U(x), \alpha_{x,u}, p(\cdot \mid x, u), g(x, u)\},
\]

where \( g(x, u) \) denotes the expected cost until the subsequent decision epoch, given that the system is in state \( x \) and control \( u \) is chosen at the current decision epoch. The cost \( g(x, u) \) is specified below (cf. page 83 and (4.22)). Throughout this work, we assume that the problem data is time-homogeneous, i.e., neither \( S, U(x), \alpha_{x,u}, p(x \mid x, u) \), nor \( g(x, u) \) depend on time.

The model in the application in Chapter 5 is an instance of a continuous-time MDP. Discrete-time MDPs are of particular importance, because every continuous-time MDP can be transformed to an “equivalent” discrete-time MDP, as we discuss in the following.
Auxiliary discrete-time MDP of a continuous-time MDP

In this section it is shown how a continuous-time MDP can be transformed into an “equivalent” discrete-time MDP. The latter is equivalent in the sense that e.g., an optimal policy of the (auxiliary) discrete-time MDP is also optimal for the (original) continuous-time MDP. The equivalence is discussed in Section 4.2.4 on page 99.

Let $M_{\text{cont}} = \{S, U(x), \alpha_{x,u}, p(\cdot | x,u), g(x,u)\}$ be a continuous-time MDP. The construction of the auxiliary discrete-time MDP of $M_{\text{cont}}$ is intimately related to the uniformization of continuous-time homogeneous Markov chains (cf. Section 4.1.2, page 74).

Choose $\eta \in \mathbb{R}_{>0}$ that satisfies

$$\alpha_{x,u} < \eta$$

for all $x \in S, u \in U(x)$.

We refer to $\eta$ as uniformization parameter. $\frac{1}{\eta}$ can be interpreted as smallest (over all $x \in S, u \in U(x)$) expected holding time. The discrete-time transition (jump) probabilities are defined by (cf. (4.12))

$$\tilde{p}(\bar{x} | x,u) := \begin{cases} \frac{\alpha_{x,u}}{\eta} p(\bar{x} | x,u) & \text{if } \bar{x} \neq x \\ 1 - \frac{\alpha_{x,u}}{\eta} & \text{if } \bar{x} = x \end{cases}$$

for all $x, \bar{x} \in S, u \in U(x)$. From the definition of $\eta$ (4.16) we have for any $x \in S, u \in U(x)$

$$\tilde{p}(x | x,u) > 0 ,$$

and for $\bar{x} \in S, \bar{x} \neq x$, we have

$$\tilde{p}(\bar{x} | x,u) > 0 \iff p(\bar{x} | x,u) > 0 .$$

The discrete-time cost is defined by

$$\tilde{g}(x,u) := \alpha_{x,u} g(x,u)$$

for $x \in S$ and $u \in U(x)$, i.e., $\tilde{g}(x,u)$ is the expected cost rate until the next decision epoch (cost per time unit).

---

75 cf. [Put94, p. 568].
76 cf. [Put94, p. 556].
77 cf. [Ber07, p. 315].
78 cf. [Put94, p. 568].
4.2 On Markov decision processes

The auxiliary discrete-time Markov decision process is denoted by \( \tilde{M}_\eta^{\text{cont}} \) and is characterized by

\[
\tilde{M}_\eta^{\text{cont}} = \{S, U(x), \tilde{p}(\cdot \mid x, u), \tilde{g}(x, u)\}.
\]

(4.21)

\( \tilde{M}_\eta^{\text{cont}} \) is referred to as uniformization of \( M^{\text{cont}} \) (with uniformization parameter \( \eta \)).

Cost and optimization problem

We assume that the aim of the decision maker is to find a policy that is optimal with respect to a given performance criterion. Typical performance criteria consider the discounted or undiscounted expected cost over a finite or an infinite time horizon. Motivated by the application in Chapter 5, we restrict our attention to undiscounted cost over an infinite time horizon; more precisely, we consider the average expected cost over an infinite time horizon. There is, of course, a whole theory on finite horizon and discounted problems as well, but we do not summarize those results here.\(^{79}\)

Our definition of the average expected cost of a policy (cf. (4.23) below) is based on the expected cost between decision epochs. We assume the following (simple) cost model which suits the application in Chapter 5. Being in state \( x \in S \) and choosing control \( u \in U(x) \) at the current decision epoch, a cost at rate \( \hat{g}(x, u) \) is accumulated until the next decision epoch. The expected cost until the subsequent decision epoch\(^{80}\) \( g(x, u) \) is thus given by\(^{81}\)

\[
g(x, u) = \begin{cases} 
\hat{g}(x, u) \cdot \Delta & \text{for a discrete-time MDP } M^{\text{discr}} \\
\hat{g}(x, u) \cdot \frac{1}{\alpha_{x,u}} & \text{for a continuous-time MDP } M^{\text{cont}}.
\end{cases}
\]

(4.22)

Consequently, the cost of the uniformization of a continuous-time MDP is given by

\[
\tilde{g}(x, u) = \hat{g}(x, u).
\]

Remark 4.2.2 In a more general setting, the cost might additionally consist of a fixed cost (at decision epochs) that is independent of the

\(^{79}\) The interested reader is referred to the literature, e.g. [Put94, Ber05, Ber07].

\(^{80}\) cf. [Put94, p. 567].

\(^{81}\) with the convention that \( \frac{1}{0} = \infty \).
length of the sojourn in the corresponding state. Moreover, the cost rate may depend on the subsequent state (cf. the cost model in [Ber07, p. 310] for semi-Markov decision processes, a generalization of continuous-time MDPs, cf. Remark 4.2.4). In the cost model in [Put94, Section 11.1.2] (also for semi-Markov decision processes), the cost rate depends on the state of the natural process which is allowed to change between decision epochs. However, by adapting the expected cost \( g(x, u) \) accordingly, the subsequent theory also applies to these more general cost models.

The average expected cost\(^84\) \( J_\pi(x_0) \) of policy \( \pi = \{\mu_0, \mu_1, \ldots\} \in \Pi \) starting from the initial state \( x_0 \in S \) is defined as

\[
J_\pi(x_0) := \lim_{N \to \infty} \sup \frac{\mathbb{E} \left[ \sum_{n=0}^{N-1} g(X_{\pi,n}, Y_{\pi,n}) \mid X_{\pi,0} = x_0, \pi \right]}{\mathbb{E} \left[ \sum_{n=0}^{N-1} \tau_{n+1} - \tau_n \mid X_{\pi,0} = x_0, \pi \right]}, \tag{4.23}
\]

where the random variable \( Y_{\pi,n}, n \in \mathbb{N}_0 \) represents the control that is chosen at decision epoch \( \tau_n \) under policy \( \pi \). For a policy \( \pi = \{\mu_0, \mu_1, \ldots\} \in \Pi^{\text{MD}} \) with deterministic Markovian decision rules \( \mu_n \), we have \( Y_{\pi,n} = \mu_n(X_{\pi,n}) \) for all \( n \in \mathbb{N}_0 \).

The optimal average expected cost\(^85\) starting from the initial state \( x_0 \in S \) is defined by the following optimization problem

\[
J^*(x_0) := \inf_{\pi \in \Pi} J_\pi(x_0). \tag{4.24}
\]

[Put94] considers rewards instead of costs and thus uses “\( \lim \inf \)” in (4.23) and “\( \sup \)” in (4.24).\(^{86}\) [Ber07] also considers costs.

For finite \( S \) and \( U \) and unichain MDPs (where the embedded Markov chain corresponding to every deterministic stationary policy is unichain) there exists a deterministic stationary policy that is optimal (cf. Theorems 4.2.10 and 4.2.17 below).

\(82\) cf. [Put94, Section 11.1.2].

\(83\) cf. [Ber07, p. 310], [Put94, p. 549].

\(84\) cf. [Put94, p. 548].

\(85\) cf. [Put94, p. 549].

\(86\) This also affects some of the subsequent results, where we use “\( \min \)” instead of “\( \max \)” as in [Put94].
Remark 4.2.3 For a continuous-time MDP, another (maybe more natural) approach to define the average expected cost of a policy \( \pi \) is

\[
\bar{J}_\pi(x_0) := \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}\left[ \int_0^T \hat{g}(X_{\pi,t}^{\text{nat}}, \bar{Y}_{\pi,t}) \, dt \mid X_{\pi,0} = x_0, \pi \right],
\]

where \( \bar{Y}_{\pi,t} \) denotes the control that is applied at time \( t \) under policy \( \pi \) (\( \bar{Y}_{\pi,t} \) is only allowed to change at decision epochs).\(^{87,88}\) However, it can be shown (for finite \( S \) and \( U \) and a unichain continuous-time MDP with \( \alpha_{x,u} \neq 0 \) for all \( x \in S, u \in U(x) \)) that the optimal average expected cost is equal for both definitions and that an optimal deterministic stationary policy\(^{89}\) for (4.24) is also optimal for the alternative problem formulation.\(^{90}\)

Remark 4.2.4 For models in continuous-time, the assumption that the holding times are exponentially distributed (as for continuous-time MDPs) might be too restricting. In fact, semi-Markov decision processes (semi-MDPs) consider holding times with arbitrary probability distributions. A semi-MDP can also be transformed to an equivalent auxiliary discrete-time MDP (similar to (4.21)); the transformation is based on the expected holding times. The optimal average expected cost and an optimal (deterministic stationary) policy can be found via the auxiliary discrete-time MDP. For the theory of semi-MDPs, we refer to the literature, e.g., Chapter 11 in [Put94].

4.2.2 Discrete-time MDPs with average expected cost

In this section, we summarize the theoretical basis to solve (4.24) for a discrete-time Markov decision process \( \mathcal{M}^{\text{discr}} = \{S, U(x), p(\cdot \mid x, u), g(x, u)\} \) as defined in Section 4.2.1. Solution methods (for the special case of unichain discrete-time MDPs) are discussed in Section 4.2.3. Recall our assumption that the set of states \( S \) and the set of controls \( U \) are finite. Several of the subsequent results hold in greater generality than they are stated here. For more general results, we refer to the literature, e.g., Chapters 8 and 9 in [Put94].

\(^{87}\) cf. [Put94, p. 548].
\(^{88}\) cf. also Theorem 4.1.10 on CTHMCs.
\(^{89}\) which exists, cf. Section 4.2.4.
\(^{90}\) See [Put94, Corollary 11.4.8] and the references therein.
By changing the time unit, we assume in the following without loss of
generality that (cf. (4.13))

$$\Delta = 1.$$ 

In the discrete setting, the optimal average expected cost starting from
the initial state $x_0 \in S$ is given by the optimization problem (cf. (4.24)
and (4.23))

$$J^*(x_0) = \inf_{\pi \in \Pi} J_\pi(x_0)$$

$$= \inf_{\pi \in \Pi} \left\{ \limsup_{N \to \infty} \frac{1}{N} \cdot \mathbb{E} \left[ \sum_{n=0}^{N-1} g(X_{\pi,n}, Y_{\pi,n}) \mid X_{\pi,0} = x_0, \pi \right] \right\}. \quad (4.25)$$

For finite $S$ and $U$ it can be shown that there exists a deterministic sta-
tionary policy that is optimal for (4.25), i.e., there is a deterministic
Markovian decision rule that, when repeatedly applied, leads to the op-
timal average expected cost. In the subsequent discussion, we thus focus
on deterministic Markovian decision rules.

We use the following notations. For a deterministic Markovian decision
rule $\mu \in D^{MD}$, let $g_\mu \in \mathbb{R}^{|S|}$ denote the vector of its expected costs, i.e.,

$$g_\mu(x) := g(x, \mu(x)) \quad x \in S.$$ 

Also, we define the stochastic matrix $P_\mu \in \mathbb{R}^{|S| \times |S|}$ element-wise by

$$(P_\mu)_{x,\bar{x}} := p(\bar{x} \mid x, \mu(x)) \quad x, \bar{x} \in S,$$

i.e., the embedded Markov chain corresponding to the deterministic sta-
tionary policy $\{\mu, \mu, \ldots\}$ has transition matrix $P_\mu$. We further use the
notations $J_\pi$ (for $\pi \in \Pi$) and $J^*$ for the vectors in $\mathbb{R}^{|S|}$ that con-
tain $J_\pi(x_0), x_0 \in S$, and $J^*(x_0), x_0 \in S$, respectively. Finally, let
$e := (1, \ldots, 1)^T \in \mathbb{R}^{|S|}$ denote the summation vector.

**Remark 4.2.5** Let $\pi = \{\mu_0, \mu_1, \ldots\} \in \Pi^{MD}$ (i.e., each $\mu_n, n \in \mathbb{N}_0$, is
a deterministic Markovian decision rule). Define the $n$-step transition

---

91 [Put94, Theorem 9.1.8].
92 cf. [Ber07, p. 175], [Put94, p. 137].
matrix\(^{93}\) \(P^{(n)}_\pi \in \mathbb{R}^{\left|S\right| \times \left|S\right|}\), \(n \in \mathbb{N}_0\), of \(\pi\) by

\[
P^{(0)}_\pi := I \\
P^{(n)}_\pi := P_{\mu_{n-1}} \cdots P_{\mu_0} \quad n \in \mathbb{N}
\]

where \(I \in \mathbb{R}^{\left|S\right| \times \left|S\right|}\) denotes the identity matrix, i.e., \((P^{(n)}_\pi)_{x,\bar{x}} = \mathbb{P}[X_{\pi,n} = \bar{x} \mid X_{\pi,0} = x, \pi]\). Then, \(J_\pi\) (cf. (4.25)) can be written as\(^ {94}\)

\[
J_\pi = \lim \sup_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} P^{(n)}_\pi g_{\mu_n}.
\]

The following theorem shows that for deterministic stationary policies, the limit in (4.25) (and in (4.26)) exists.\(^ {95}\)

**Theorem 4.2.6** Let \(S\) be finite. Let \(\mu \in \Pi_{MD}^{MD}\) be a deterministic stationary policy and let \(P^*_\mu := \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} P^n_\mu\) denote the limiting matrix\(^ {96}\) of \(P_\mu\). Then the limit in (4.25) (and (4.26)) corresponding to the deterministic stationary policy \(\mu\) exists and

\[
J_\mu = P^*_\mu g_\mu.
\]

The previous theorem generalizes the result of Theorem 4.1.7 to non-unichain Markov chains where the average expected cost may depend on the initial state.

**Equal (optimal) average expected cost for all initial states, classification of discrete-time MDPs**

For discrete-time MDPs with average expected cost, the probabilistic structure of the underlying discrete-time homogeneous Markov chains, e.g., number of recurrent communicating classes, is important both for analysis and algorithms.\(^ {97}\) In the sequel we discuss structural assumptions for discrete-time MDPs which assure that the optimal average expected cost \(J^*(x_0)\) is equal for all initial states \(x_0 \in S\). In this case, we

---

\(^{93}\) cf. [Put94, p. 138].

\(^{94}\) cf. [Put94, p. 332].

\(^{95}\) [Put94, Proposition 8.1.1]. In fact, the result is proven for randomized stationary policies (with appropriate modification of definitions).

\(^{96}\) cf. (4.7).

\(^{97}\) cf. [Ber07, p. 174].
can write

$$J^* = \lambda^* e$$  \hspace{1cm} (4.27)

for some $\lambda^* \in \mathbb{R}$. We start with some properties of the average expected cost of a deterministic stationary policy.\textsuperscript{98}

**Theorem 4.2.7** Let $S$ be finite and let $\mu \in \Pi^{MD}$ be a deterministic stationary policy. Assume that the embedded Markov chain $X_\mu$ corresponding to $\mu$ is unichain. Then:

(i) The average expected cost $J_\mu(x_0)$ of $\mu$ is equal for all initial states $x_0 \in S$, i.e.,

$$J_\mu = \lambda_\mu e$$

for some $\lambda_\mu \in \mathbb{R}$.

(ii) The average expected cost $\lambda_\mu$ in (i) is uniquely determined by the following system of equations with variables $\lambda_\mu \in \mathbb{R}$ and $h \in \mathbb{R}^{|S|}$

$$\lambda_\mu e + h = g_\mu + P_\mu h.$$  \hspace{1cm} (4.28)

(iii) The average expected cost $\lambda_\mu$ in (i) is given by

$$\lambda_\mu = \nu_\mu^T g_\mu,$$

where $\nu_\mu$ denotes the unique stationary distribution of $X_\mu$ (which is given by any of the identical rows of $P_{\mu}^*$, the limiting matrix of $P_\mu$).

Whereas $\lambda_\mu$ in the above theorem is uniquely determined by (4.28), $h$ is only determined up to an additive constant.\textsuperscript{99}

The following classification of discrete-time MDPs (adapted from [Put94, p. 348]) is based on the “reachability” of states. For the notions related to discrete-time homogeneous Markov chains, we refer to Section 4.1.1.

\textsuperscript{98} The theorem also holds for randomized stationary polices (with appropriate modification of definitions). For (i), see [Put94, Proposition 8.2.1]. For (ii), see [Put94, Corollary 8.2.7]. (iii) follows from Theorems 4.1.6 and 4.2.6.

\textsuperscript{99} [Put94, Corollary 8.2.7].
A discrete-time MDP (with finite $S$ and $U$) is called

- **recurrent** if the embedded Markov chain corresponding to every deterministic stationary policy is irreducible (and thus the corresponding communicating class is recurrent).
- **unichain** if the embedded Markov chain corresponding to every deterministic stationary policy is unichain.
- **weakly communicating**\(^{100}\) if the state space $S$ can be partitioned into a set of states $S_{\text{access}}$ and a (possibly empty) set of states $S_{\text{transient}}$ such that:
  
  (a) For all $x, \bar{x} \in S_{\text{access}}$, $\bar{x}$ is *accessible*\(^{101}\) from $x$, i.e., there exists a deterministic stationary policy $\mu$ and an integer $n \geq 1$ such that
  
  $$ \mathbb{P}[X_{\mu,n} = \bar{x} \mid X_{\mu,0} = x, \mu] > 0. $$

  (b) The states in $S_{\text{transient}}$ are transient for every policy.\(^{102}\)
- **communicating** if it is weakly communicating with $S_{\text{transient}} = \emptyset$.
- **multichain** if the embedded Markov chain corresponding to at least one stationary policy has at least two (closed) recurrent communicating classes.

From the definitions, it is immediately clear that a recurrent MDP is unichain and communicating and that a communicating MDP is weakly communicating. It also follows that:\(^{103}\)

**Theorem 4.2.8** A unichain discrete-time MDP is weakly communicating.

---

\(^{100}\) cf. [Ber07, Definition 4.2.2]. In [Ber07], a weakly communicating MDP is referred to as one that satisfies the *weak accessibility* condition.

\(^{101}\) cf. [Ber07, Definition 4.2.1].

\(^{102}\) $S_{\text{access}}$ is *closed* (cf. [Put94, p. 348, 588]), i.e., no state in $S_{\text{transient}} = S \setminus S_{\text{access}}$ is accessible from any state in $S_{\text{access}}$: suppose there exist $x \in S_{\text{access}}$, $\bar{x} \in S_{\text{transient}}$, $\mu \in \Pi_{\text{MD}}$, and $n \geq 1$ such that $\mathbb{P}[X_{\mu,n} = \bar{x} \mid X_{\mu,0} = x, \mu] > 0$. As $S$ is finite, there is $\hat{x} \in S$ which is recurrent for $\mathcal{X}_{\mu}$ (thus $\hat{x} \in S_{\text{access}}$) and $\hat{n} \geq 1$ such that $\mathbb{P}[X_{\mu,\hat{n}} = \hat{x} \mid X_{\mu,0} = \bar{x}, \mu] > 0$ (the set of transient states is left after a finite number of steps). Consequently, $\bar{x}$ belongs to $S_{\text{access}}$, which is a contradiction.

\(^{103}\) [Ber07, Proposition 4.2.5].
[Tsi07] shows that checking whether a discrete-time MDP (with finite \( S \) and \( U \)) is unichain is NP-hard.\(^\text{104}\) If there exists a state \( x \in S \) that is either (i) recurrent for the embedded Markov chain corresponding to each deterministic stationary policy or (ii) stopping, i.e., \( p(x \mid x, u) = 1 \) for some \( u \in U(x) \), [FY08] provide polynomial algorithms (in \( |S| \) and \( \sum_{x \in S} |U(x)| \)) for checking the unichain condition. They also provide a polynomial algorithm to detect whether there is a state of type (i). Clearly, states of type (ii) can also be identified with a polynomial algorithm. Moreover, deciding whether a discrete-time MDP (with finite \( S \) and \( U \)) is recurrent, communicating, or weakly communicating is also polynomially solvable.\(^\text{105}\)

An important consequence of the above structural classification on the optimal average expected cost is given in the following theorem.\(^\text{106}\)

**Theorem 4.2.9** Let \( S \) and \( U \) be finite and assume a discrete-time MDP that is weakly communicating. Then the optimal average expected cost is independent of the initial state, i.e., (4.27) holds.

The subsequent Theorem 4.2.10 establishes this result for the special case of a unichain discrete-time MDP.

**Review of optimality equations and existence of optimal policies in the unichain case**

Let \( \mathcal{M}^{\text{discr}} = \{ S, U(x), p(\cdot \mid x, u), g(x, u) \} \) denote a discrete-time Markov decision process which is assumed to be unichain. The following theorem characterizes the optimal average expected cost and a corresponding optimal policy in terms of optimality equations.\(^\text{107}\)

\(^{104}\) More precisely, he proves that deciding whether there exists a deterministic stationary policy such that the corresponding embedded Markov chain has at least two recurrent communicating classes is NP-complete, from which the NP-hardness follows, cf. [GJ79, p. 114].

\(^{105}\) See [Kal02, Algorithms 2, 3, and 4, respectively] and the references therein. Note that by [Put94, Proposition 8.3.1], the definition of weakly communicating in [Kal02] is equivalent to the definition given in this thesis. For deciding whether the MDP is communicating, weakly communicating, or multichain, cf. also [Put94, p. 351].

\(^{106}\) [Put94, Theorem 8.3.2 and Theorem 9.1.8].

\(^{107}\) See [Put94, Theorem 8.4.1(c), Theorem 8.4.3(a), and Theorem 8.4.4] for (i), (ii), and (iii), respectively.
Theorem 4.2.10 Let $S$ and $U$ be finite and assume $\mathcal{M}^{\text{discr}}$ is unichain.

(i) If $\lambda^* \in \mathbb{R}$ and $h^* \in \mathbb{R}^{|S|}$ satisfy

$$\lambda^* + h^*(x) = \min_{u \in U(x)} \left[ g(x, u) + \sum_{\bar{x} \in S} p(\bar{x} \mid x, u) h^*(\bar{x}) \right], \quad x \in S,$$

(4.29) then $J^* = \lambda^* e$ is the optimal average expected cost.

(ii) There exist $\lambda^* \in \mathbb{R}$ and $h^* \in \mathbb{R}^{|S|}$ that satisfy (4.29).

(iii) If $\lambda^* \in \mathbb{R}$ and $h^* \in \mathbb{R}^{|S|}$ satisfy (4.29) and $\mu^* \in D^{MD}$ is such that $\mu^*(x)$ attains the minimum in (4.29) for all $x \in S$, then the deterministic stationary policy $\{\mu^*, \mu^*, \ldots\}$ is optimal for (4.25).

For finite $S$ and $U$ and a unichain discrete-time MDP, Theorem 4.2.10 establishes the existence of an optimal deterministic stationary policy and the optimization problem (4.25) can be stated as

$$J^*(x_0) = \min_{\mu \in \Pi^{MD}} \left\{ \lim_{N \to \infty} \frac{1}{N} \cdot \mathbb{E}\left[ \sum_{n=0}^{N-1} g(X_{\pi,n}, \mu(X_{\pi,n})) \mid X_{\pi,0} = x_0, \pi \right] \right\}. \quad (4.30)$$

The minimum is attained because there are only finitely many deterministic stationary policies and the limit exists due to Theorem 4.2.6. Moreover, the optimal average expected cost is equal for all initial states, i.e., (4.27) holds.

4.2.3 Solution methods for unichain discrete-time MDPs

There are three main solution methods for discrete-time MDPs with average expected cost: value iteration, policy iteration, and linear programming. In this section, we summarize the properties of the three methods and discuss their advantages and drawbacks. We argue that, for very large state spaces (e.g., as in the application in Chapter 5), value iteration seems to be the most suitable method.

As before, we assume finite $S$ and $U$ and we assume that the discrete-time MDP $\mathcal{M}^{\text{discr}} = \{S, U(x), p(\cdot \mid x, u), g(x, u)\}$ is unichain, which assures (4.27) and guarantees a solution to (4.29) and the existence of an optimal deterministic stationary policy.
As mentioned in [ZES09], the only proof that discrete-time MDPs (with finite $S$ and $U$) with the average expected cost criterion can be solved in polynomial time relies on the formulation as an LP. The authors provide a strongly polynomial algorithm for a special class of recurrent MDPs. In 1987, [PT87] proved that finding an optimal deterministic stationary policy for a discrete-time MDP (with finite $S$ and $U$) is P-complete\textsuperscript{108}.

**Value iteration for unichain discrete-time MDPs**

The basic idea of value iteration is to use dynamic programming, starting with some arbitrary terminal cost, to determine the optimal (total) expected cost over $N$ steps (i.e., up to decision epoch $\tau_N$).\textsuperscript{109} Starting from an (arbitrary) vector $h_0^{\text{pure}} \in \mathbb{R}^{|S|}$, the (pure) value iteration\textsuperscript{110} is defined by

\[
h^{n+1}_{\text{pure}} := \min_{\mu \in D_{MD}} [g_{\mu} + P_{\mu} h^n_{\text{pure}}] \quad n \in \mathbb{N}_0,
\]

i.e., in componentwise notation

\[
h^{n+1}_{\text{pure}}(x) := \min_{u \in U(x)} [g(x, u) + \sum_{\bar{x} \in S} p(\bar{x} | x, u) h^n_{\text{pure}}(\bar{x})] \quad x \in S.
\]

The following theorem shows that the differences $h^{n+1}_{\text{pure}}(x) - h^n_{\text{pure}}(x)$ allow to calculate bounds on the optimal average expected cost $\lambda^*$ (where $J^* = \lambda^* e$ according to Theorem 4.2.10).\textsuperscript{111}

**Theorem 4.2.11** Let $S$ and $U$ be finite and assume $\mathcal{M}_{\text{discr}}$ is unichain. For $h \in \mathbb{R}^{|S|}$, let

\[
h' := \min_{\mu \in D_{MD}} [g_{\mu} + P_{\mu} h]
\]

and

\[
\mu' \in \arg\min_{\mu \in D_{MD}} [g_{\mu} + P_{\mu} h].
\]

\textsuperscript{108} A problem is P-complete if it is in P and every problem in P is reducible to it in logarithmic space, cf. [PT87].

\textsuperscript{109} cf. [Ber07, p. 204].

\textsuperscript{110} cf. [Put94, p. 364].

\textsuperscript{111} [Put94, Theorem 8.5.5].
4.2 On Markov decision processes

Let \( \lambda_{\mu'} \) denote the average expected cost of the deterministic stationary policy \( \{\mu', \mu', \ldots\} \).\(^{112}\) Then

\[
\min_{x \in S} [h'(x) - h(x)] \leq \lambda^* \leq \max_{x \in S} [h'(x) - h(x)].
\]

There exist different conditions which assure that the differences \( h^{n+1}_{\text{pure}}(x) - h^n_{\text{pure}}(x) \) converge to a single value (and thus, by Theorem 4.2.11, yield the optimal average expected cost).\(^{113}\) However, as some components of \( h^{n+1}_{\text{pure}} \) may diverge to \( \pm \infty \), the direct use of (pure) value iteration may cause numerical difficulties.\(^{114}\) To avoid this, a (per iteration fixed) scalar value is subtracted from all components of \( h^{n+1}_{\text{pure}} \). This is known as relative value iteration. The corresponding method is the relative value iteration algorithm\(^{115,116}\) that is defined as follows:

1. Specify the stopping accuracy \( \epsilon > 0 \), choose a reference state \( x_r \in S \), choose an (arbitrary) vector \( \hat{h}^0_{\text{rel}} \in \mathbb{R}^{|S|} \), and set \( h^0_{\text{rel}} := \hat{h}^0_{\text{rel}} - \hat{h}^0_{\text{rel}}(x_r)e \), and set \( n := 0 \).

2. Calculate \( \hat{h}^n_{\text{rel}} \) by

\[
\hat{h}^{n+1}_{\text{rel}} := \min_{\mu \in D_{MD}} [g_\mu + P_\mu h^n_{\text{rel}}]
\]

and set

\[
h^{n+1}_{\text{rel}} := \hat{h}^{n+1}_{\text{rel}} - \hat{h}^{n+1}_{\text{rel}}(x_r)e.
\]

3. If

\[
\max_{x \in S} [\hat{h}^{n+1}_{\text{rel}}(x) - h^n_{\text{rel}}(x)] - \min_{x \in S} [\hat{h}^{n+1}_{\text{rel}}(x) - h^n_{\text{rel}}(x)] < \epsilon
\]

go to step 4. Otherwise increment \( n \) by 1 and go to step 2.

\(^{112}\) cf. Theorem 4.2.7(i).

\(^{113}\) [Put94, Theorem 8.5.2 and Theorem 8.5.3].

\(^{114}\) cf. [Put94, p. 373], [Ber07, p. 206].

\(^{115}\) cf. [Put94, p. 373].

\(^{116}\) The algorithm in [Put94, p. 373] uses \( \hat{h}^n_{\text{rel}} \) instead of \( h^n_{\text{rel}} \) in steps 3 and 4. However, as \( \hat{h}^n_{\text{rel}} = h^n_{\text{rel}} + \hat{h}^n_{\text{rel}}(x_r)e \), the stopping criteria in step 3 and the chosen decision rules in step 4 are identical. In our formulation, there is no need to store \( \hat{h}^n_{\text{rel}}(x_r) \) (or even \( \hat{h}^n_{\text{rel}} \)). In addition, the maximum/minimum values of the stopping criterion in step 3 provide the upper/lower bounds on \( \lambda^* \) of Theorem 4.2.11.
4. Choose a decision rule $\mu_\epsilon \in D^{MD}$ that satisfies
\[
\mu_\epsilon \in \arg\min_{\mu \in D^{MD}} \{ g_\mu + P_\mu h^n_{rel} \},
\]
i.e.,
\[
\mu_\epsilon(x) \in \arg\min_{u \in U(x)} [g(x, u) + \sum_{\bar{x} \in S} p(\bar{x} | x, u) h^n_{rel}(\bar{x})] \quad x \in S \quad (4.31)
\]
and stop.

Observe that if the relative value iteration algorithm terminates, the maximum/minimum value in step 3 provide an interval of length at most $\epsilon$ that captures the optimal average expected cost $\lambda^*$ (cf. Theorem 4.2.11).

The (pure) value iteration algorithm\textsuperscript{117} is obtained from the relative value iteration algorithm by setting $h^0_{rel} := \hat{h}^0_{rel}$ in step 1 and by setting $h^{n+1}_{rel} := \hat{h}^{n+1}_{rel}$ in step 3. Due to the “renormalization” of the value iterates, the relative value iteration algorithm is numerically more stable than the (pure) value iteration algorithm.\textsuperscript{118} However, the convergence properties of both methods are identical.\textsuperscript{119}

\textsuperscript{117} cf. [Put94, p. 364].
\textsuperscript{118} cf. [Put94, p. 373].
\textsuperscript{119} cf. [Ber07, p. 208] (and [Put94, p. 373]). By induction, it follows that for each $n \in \mathbb{N}_0$, there exists a scalar $c^n \in \mathbb{R}$ such that
\[
h^n_{rel} = h^n_{pure} - c^n e, \quad (4.32)
\]
i.e., the value iterates only differ by a constant value. For a vector $v \in \mathbb{R}^{|S|}$, define $\text{span}(v) := \max_{x \in S} v(x) - \min_{x \in S} v(x)$ and observe $\text{span}(v + ce) = \text{span}(v)$ for all $c \in \mathbb{R}$ (cf. [Put94, p. 196]). By definition, $\hat{h}^{n+1}_{rel} = h^{n+1}_{rel} + h^{n+1}_{rel}(x_r)e$, i.e.,
\[
\hat{h}^{n+1}_{rel} = h^n_{pure} + (\hat{h}^{n+1}_{rel}(x_r) - c^{n+1})e
\]
and thus
\[
\text{span}(h^n_{pure} - h^n_{rel}) = \text{span}(\hat{h}^{n+1}_{rel} - h^n_{rel}), \quad (4.33)
\]
i.e., the values of the termination criterion in step 3 are identical for the pure and the relative value iteration algorithm. Also, as
\[
\arg\min_{\mu} \{ g_\mu + P_\mu h^n_{pure} \} \overset{(4.32)}{=} \arg\min_{\mu} \{ g_\mu + P_\mu (h^n_{rel} + c^n e) \}
\]
\[
= \arg\min_{\mu} \{ g_\mu + P_\mu h^n_{rel} \},
\]
both algorithms identify the same decision rules as $\epsilon$-optimal in step 4.
One sufficient condition for finite termination of the relative value iteration algorithm that suits our purposes is given in the following theorem.\textsuperscript{120}

**Theorem 4.2.12** Let $S$ and $U$ be finite. Assume that the discrete-time MDP is unichain and $p(x \mid x, u) > 0$ for all $x \in S$, $u \in U(x)$. Then for any $\epsilon > 0$, $x_r \in S$, $\hat{h}^{0}_{rel} \in \mathbb{R}^{|S|}$:

(i) The (relative) value iteration algorithm terminates after a finite number of iterations.

(ii) The deterministic stationary policy $\{\mu_\epsilon, \mu_\epsilon, \ldots\}$ with $\mu_\epsilon$ as defined by (4.31) is $\epsilon$-optimal.

**Remark 4.2.13** The condition $p(x \mid x, u) > 0$ for all $x \in S$, $u \in U(x)$ in Theorem 4.2.12 can always be obtained by a so-called aperiodicity transformation, cf. [Put94, Section 8.5.4]. For a deterministic Markovian decision rule $\mu$, the associated transition matrix $P_\mu$ is replaced by the transformed stochastic matrix $\gamma P_\mu + (1 - \gamma)I$ for a fixed $0 < \gamma < 1$. It can be shown that the optimal average expected cost as well as the sets of optimal stationary policies for the original and the transformed problem are identical.\textsuperscript{121} The aperiodicity transformation does not have to be carried out explicitly but can be incorporated in the algorithm (cf. [Ber07, p. 216]).

The main advantage of (relative) value iteration is its iterative character, i.e., the problem data can be processed sequentially (one state after the other). Also, the calculations in each iteration are rather “easy” (multiplication, addition, and comparison of values).

**Policy iteration for unichain discrete-time MDPs**

The policy iteration algorithm is an iterative procedure where each iteration consists of two steps. In the first step, a given (deterministic stationary) policy is evaluated, i.e., its average expected cost is calculated according to (4.28). Then, in the second step, an improved policy is sought.\textsuperscript{122}

\textsuperscript{120} (i): due to Theorem 4.2.10 it suffices to consider deterministic stationary policies (cf. (4.30)). The claim follows from [Put94, Theorem 8.5.3(c) and Theorem 8.5.2(b)] by using (4.33) of Footnote 119. (ii) follows from Theorem 4.2.11 (cf. [Put94, Theorem 8.5.6(a)]).

\textsuperscript{121} [Put94, Corollary 8.5.9].

\textsuperscript{122} cf. [Ber07, p. 229].
The *policy iteration algorithm*\(^{123}\) is defined as follows:

1. Choose a reference state \(x_r \in S\), set \(n := 0\), and choose an arbitrary (deterministic Markovian) decision rule \(\mu^n \in D^{MD}\).

2. **Policy evaluation step.** Determine a scalar \(\lambda^n \in \mathbb{R}\) and a vector \(h^n \in \mathbb{R}^{|S|}\) that solve\(^{124}\)

\[
\lambda^n e + h^n = g_{\mu^n} + P_{\mu^n} h^n,
\]

i.e., in componentwise notation

\[
\lambda^n + h^n(x) = g(x, \mu^n(x)) + \sum_{\bar{x} \in S} p(\bar{x} \mid x, \mu^n(x)) h^n(\bar{x}) \quad x \in S,
\]

\[
h^n(x_r) = 0.
\]

3. **Policy improvement step.** Find a (deterministic Markovian) decision rule \(\mu^{n+1} \in D^{MD}\) that satisfies

\[
\mu^{n+1} \in \arg \min_{\mu \in D^{MD}} \{g_{\mu} + P_{\mu} h^n\},
\]

i.e., in componentwise notation

\[
\mu^{n+1}(x) \in \arg \min_{u \in U(x)} \left[ g(x, u) + \sum_{\bar{x} \in S} p(\bar{x} \mid x, u) h^n(\bar{x}) \right], \quad x \in S, \quad (4.34)
\]

and set \(\mu^{n+1}(x) = \mu^n(x)\) if \(\mu^n(x)\) satisfies (4.34).

4. If \(\mu^{n+1} = \mu^n\), stop and set \(\mu^* := \mu^n\). Otherwise increment \(n\) by 1 and go to step 2.

From (4.28) we have that \(\lambda^n\) in step 2 is (indeed) the average expected cost of the deterministic stationary policy \(\{\mu^n, \mu^n, \ldots\}\), i.e., \(J_{\mu^n} = \lambda^n e\). Observe that if \(\mu^{n+1} = \mu^n\) (i.e., if the algorithm terminates), we have

\[
\min_{\mu \in D^{MD}} \{g_{\mu} + P_{\mu} h^n\} \quad (\text{step 3}) \quad g_{\mu^{n+1}} + P_{\mu^{n+1}} h^n
\]

\[
= g_{\mu^n} + P_{\mu^n} h^n \quad (\text{step 2}) \quad \lambda^n e + h^n.
\]

\(^{123}\) cf. [Put94, p. 378].

\(^{124}\) The equation \(h^n(x_r) = 0\) is used to have a unique solution of the equation system (see [Put94, Corollary 8.2.7]). There exist alternative approaches to achieve this, cf. [Put94, p. 378].
i.e., $\lambda^n$ is the optimal average expected cost and \{\mu^n, \mu^n, \ldots\} is an optimal deterministic stationary policy (cf. (i) and (iii) of Theorem 4.2.10).

For a unichain discrete-time MDP and finite $S$ and $U$ it can be shown that the above policy iteration algorithm terminates in a finite number of iterations.\textsuperscript{125}

One advantage over value iteration is that policy iteration terminates with an optimal (and not just $\epsilon$-optimal) policy. For discrete-time MDPs with a large number of states, the main disadvantage of policy iteration is that in each iteration, a system of $|S| + 1$ linear equations in $|S| + 1$ variables has to be solved.\textsuperscript{126} A remedy is to use the so-called modified policy iteration which avoids exact evaluation of a policy and determines an approximate value for the average expected cost $\lambda_{\mu^n}$. We do not further elaborate on this here; the interested reader is referred to e.g. [Put94, Section 8.7].

### Linear programming for unichain discrete-time MDPs

The optimal average expected cost $\lambda^*$ equals the optimal objective function value of the following (primal) linear program:\textsuperscript{127}

$$
\lambda^* = \max \lambda \\
\text{s.t.} \quad \lambda + h(x) \leq g(x, u) + \sum_{\bar{x} \in S} p(\bar{x} \mid x, u) h(\bar{x}) \quad x \in S, u \in U(x) \\
\lambda \in \mathbb{R} \\
h \in \mathbb{R}^{|S|}.
$$

\textsuperscript{125} See [Put94, Theorem 8.6.6], using that (due to Theorem 4.2.10) it suffices to consider deterministic stationary policies (cf. (4.30)).

\textsuperscript{126} cf. [Ber07, p. 40] in the context of discounted problems.

\textsuperscript{127} cf. [Ber07, Section 4.5] or [Put94, Section 8.8].
The corresponding dual linear program is\(^{128}\)

\[
\min \sum_{x \in S} \sum_{u \in U(x)} g(x, u)q(x, u)
\]
\[
\text{s.t. } \sum_{u \in U(\bar{x})} q(\bar{x}, u) = \sum_{x \in S} \sum_{u \in U(x)} p(\bar{x} \mid x, u)q(x, u) \quad \bar{x} \in S
\]
\[
\sum_{x \in S} \sum_{u \in U(x)} q(x, u) = 1
\]
\[
q(x, u) \geq 0 \quad x \in S, u \in U(x).
\]

There is an intimate relation between (basic) dual feasible solutions and (randomized/deterministic) stationary policies and the stationary distributions of the corresponding embedded Markov chains. We do not elaborate on this here; the interested reader is referred to [Put94, Section 8.8].

An optimal deterministic stationary policy can be obtained from a dual optimal solution as follows.\(^{129}\) Let \(q^*(x, u), x \in S, u \in U(x),\) denote a dual optimal solution and let \(S_{q^*} := \{x \in S \mid \sum_{u \in U(x)} q^*(x, u) > 0\}\).

From the second equality constraint of the dual LP it follows that \(S_{q^*} \neq \emptyset\).

For \(x \in S_{q^*}\) define the set of controls \(U_{q^*}(x) = \{u \in U(x) \mid q^*(x, u) > 0\}\). As \(S_{q^*} \neq \emptyset, U_{q^*}(x) \neq \emptyset\) for all \(x \in S_{q^*}\). Then the (deterministic stationary) policies \(\{\mu_{q^*}, \mu_{q^*}, \ldots\}\) with \(\mu_{q^*}\) defined by

\[
\mu_{q^*}(x) = \begin{cases} 
\text{any } u \in U_{q^*}(x) & \text{if } x \in S_{q^*}, \\
\text{any } u \in U(x) & \text{for } x \notin S_{q^*}
\end{cases}
\]

are optimal.\(^{130}\)

**Remark 4.2.14** In the more general case, i.e., without assuming that the discrete-time MDP is unichain or that the optimal average expected cost is equal for all initial states, it can be shown that the optimal objective function value \(\lambda^*\) of the primal LP (which equals the optimal objective function value of the dual LP) yields the lowest optimal average expected cost over all states.\(^{131}\)

\(^{128}\) cf. [Ber07, p. 240], [Put94, p. 391].\(^{129}\) cf. [Ber07, p. 241].\(^{130}\) cf. [Ber07, p. 241, 242].\(^{131}\) cf. [Ber07, Section 4.5 and Exercise 4.16].
An advantage of the linear programming formulation is its “closed form” formulation. Also, the relations between (basic) feasible solutions of the dual LP and stationary policies may be useful. However, it is exactly the closed form (with \( \sum_{x \in S} |U(x)| \) constraints in \(|S| + 1\) variables) which may prohibit using the LP formulation for large problem instances.

Summary and comments

Certainly, for all solution methods described in this section, the problem data (i.e., the transition (jump) probabilities \( p(\tilde{x} \mid x, u) \)) and the expected cost \( g(x, u) \) have to be stored. As already mentioned, the main advantage of (relative) value iteration for large problem instances is that it processes the problem data sequentially. In each iteration, the states can be considered one after the other with rather “easy” calculations. On the contrary, policy iteration and the LP formulation need to process the data “as a whole”; policy iteration treats all states simultaneously by solving a single linear system of equations in each iteration and the LP has to consider all constraints (each corresponding to a state and a control) simultaneously. Thus, for the model in Chapter 5, due to the large number of states, iterative procedures (with comparably “cheap” iterations) seem to be most promising. In his Master’s thesis, [Zür09] solved the dual linear program with a subgradient method. From his empirical results, however, he concludes that relative value iteration allows the solution of larger problem instances than the subgradient method applied to the dual LP.

4.2.4 Solving unichain continuous-time MDPs with average expected cost

In this section the optimal average expected cost of a continuous-time MDP is characterized and it is shown that a continuous-time MDP and its auxiliary discrete-time MDPs are equivalent in the sense that they have the same optimal average expected cost and the same set of optimal (deterministic stationary) policies. Recall our assumption of a finite state space \( S \) and a finite set of controls \( U \). Some of the subsequent results hold in greater generality than they are stated here. For more general statements, we refer to the literature, e.g., [Put94, Section 11.4].

Let \( \mathcal{M}^{\text{cont}} = \{S, U(x), \alpha_{x,u}, p(\cdot \mid x, u), g(x, u)\} \) be a continuous-time
MDP as defined in Section 4.2.1. Similar to the discrete-time case (cf. Theorem 4.2.7(i)) we have:

**Theorem 4.2.15** Let $S$ be finite. Let $\mu \in \Pi$ be a stationary policy and assume that the embedded Markov chain $X_{\mu}$ corresponding to $\mu$ is unichain. Then the average expected cost $J_{\mu}(x_0)$ of $\mu$ (cf. (4.23)) is equal for all initial states $x_0 \in S$, i.e., $J_{\mu} = \lambda_{\mu}e$ for some $\lambda_{\mu} \in \mathbb{R}$.

Consistent with the definition for discrete-time MDPs, we call the continuous-time MDP $\mathcal{M}^{\text{cont}}$ unichain if the embedded Markov chain corresponding to every deterministic stationary policy is unichain.

The following theorem provides an optimality equation for continuous-time MDPs.

**Theorem 4.2.16** Let $S$ and $U$ be finite. If $\lambda^* \in \mathbb{R}$ and $h^* \in \mathbb{R}^{|S|}$ satisfy

$$h^*(x) = \min_{u \in U(x)} \left[ g(x, u) - \frac{\lambda^*}{\alpha_{x,u}} + \sum \limits_{\bar{x} \in S} p(\bar{x} \mid x, u) h^*(\bar{x}) \right] \quad x \in S, \quad (4.35)$$

then $J^* = \lambda^*e$ is the optimal average expected cost.

For unichain continuous-time MDPs, a solution to (4.35) exists (cf. (ii) of the subsequent Theorem 4.2.17).

Let $\tilde{\mathcal{M}}^{\text{cont}}_{\eta}$ denote a uniformization of $\mathcal{M}^{\text{cont}}$ with uniformization parameter $\eta$ (cf. (4.21)). As $S$ is finite, we conclude from Theorem 4.1.11:

$$\mathcal{M}^{\text{cont}} \text{ is unichain } \Leftrightarrow \tilde{\mathcal{M}}^{\text{cont}}_{\eta} \text{ is unichain.} \quad (4.36)$$

The following theorem relates the continuous-time MDP $\mathcal{M}^{\text{cont}}$ to the auxiliary discrete-time MDP $\tilde{\mathcal{M}}^{\text{cont}}_{\eta}$.\[\text{136}\]

---

132 [Put94, Proposition 11.4.1].

133 [Put94, Theorem 11.4.4].

134 Let $\mu \in \Pi^{\text{MDP}}$ be a deterministic stationary policy. The embedded Markov chain $X_{\mu}^{\eta}$ of $\mathcal{M}^{\text{cont}}_{\eta}$ corresponding to $\mu$ is an auxiliary DTHMC (as defined in Section 4.1.2, page 74) of the CTHMC $X_{\text{nat}}^{\mu}$ whose jump chain is the embedded Markov chain $X_{\mu}$ (of $\mathcal{M}^{\text{cont}}$) corresponding to $\mu$.

135 cf. also [Ber07, p. 316].

136 For (i), (iii), see [Put94, Proposition 11.4.5(a),(b)]. (ii) follows from Theorem 4.2.10(ii) and part (i) (cf. also [Put94, Theorem 11.4.6]). (iv) follows from Theorem 4.2.10(i)-(ii), part (i), and Theorem 4.2.16 (cf. also [Put94, Proposition 11.4.5(c)]).
Theorem 4.2.17 Let $S$ and $U$ be finite and assume that $\mathcal{M}^\text{cont}$ is unichain.

(i) If $\tilde{\lambda} \in \mathbb{R}$, $\tilde{h} \in \mathbb{R}^{|S|}$ satisfy the discrete-time optimality equations (cf. (4.29))

$$
\tilde{\lambda} + \tilde{h}(x) = \min_{u \in U(x)} \left[ \tilde{g}(x, u) + \sum_{\bar{x} \in S} \tilde{p}(\bar{x} \mid x, u) \tilde{h}(\bar{x}) \right] \quad x \in S,
$$

then $(\lambda^*, h^*)$ with $\lambda^* := \tilde{\lambda}$, $h^* := \frac{1}{\eta} \tilde{h}$, satisfy (4.35).

(ii) There exist $\lambda^* \in \mathbb{R}$ and $h^* \in \mathbb{R}^{|S|}$ that satisfy (4.35).

(iii) Let $\mu \in \Pi^\text{MD}$ be a deterministic stationary policy. Then, the average expected cost $J^{\mu} = \lambda^{\mu} e$ of $\mu$ for $\mathcal{M}^\text{cont}$ and the average expected cost $\tilde{J}^{\mu} = \tilde{\lambda}^{\mu} e$ of $\mu$ for $\tilde{\mathcal{M}}^\text{cont}$ are equal, i.e.,

$$
\lambda^{\mu} = \tilde{\lambda}^{\mu}.
$$

(iv) The optimal average expected cost $J^*_{\mathcal{M}^\text{cont}}$ and $J^*_{\tilde{\mathcal{M}}^\text{cont}}$ of $\mathcal{M}^\text{cont}$ and $\tilde{\mathcal{M}}^\text{cont}$, respectively, satisfy $J^*_{\mathcal{M}^\text{cont}} = \lambda^{*_{\mathcal{M}^\text{cont}}} e$ and $J^*_{\tilde{\mathcal{M}}^\text{cont}} = \lambda^{*_{\tilde{\mathcal{M}}^\text{cont}}} e$ for some $\lambda^{*_{\mathcal{M}^\text{cont}}}, \lambda^{*_{\tilde{\mathcal{M}}^\text{cont}}} \in \mathbb{R}$ and

$$
\lambda^{*_{\mathcal{M}^\text{cont}}} = \lambda^{*_{\tilde{\mathcal{M}}^\text{cont}}}.
$$

Consequently, the optimal average expected cost and an optimal policy for the (unichain) continuous-time MDP can be obtained by solving the (unichain) auxiliary discrete-time MDP. More precisely, by applying the methods of Section 4.2.3 to $\tilde{\mathcal{M}}^\text{cont}_\eta$ we can identify an $(\epsilon)$-optimal deterministic stationary policy $\tilde{\mu}^*$ (which exists by Theorem 4.2.10) and find the $(\epsilon)$-optimal average expected cost $\lambda^{*_{\tilde{\mathcal{M}}^\text{cont}_\eta}}$ of $\tilde{\mathcal{M}}^\text{cont}_\eta$. By the above Theorem 4.2.17, $\tilde{\mu}^*$ is also an $(\epsilon)$-optimal deterministic stationary policy and $\lambda^{*_{\tilde{\mathcal{M}}^\text{cont}_\eta}}$ the $(\epsilon)$-optimal average expected cost for the continuous-time MDP $\mathcal{M}^\text{cont}_\eta$.

---

137 $J^{\mu}$ and $\tilde{J}^{\mu}$ are indeed equal for all initial states, cf. Theorem 4.2.15 and Theorem 4.2.7(i), respectively.
Chapter 5

Resource-constrained power grid model (focus II)
In this chapter, we present the second model, whose focus lies on the “vulnerability” of redundantly operated power grids over time (focus II, cf. Section 2.2.1). The power grid is modeled by independent components with exponential failure and repair rates. However, a failed component can only be repaired if an available resource is assigned to this component. The model is formulated as a continuous-time Markov decision process. For a given number of resources (1st decision level, cf. Section 2.2), the optimal assignment of these resources to failed components (2nd decision level) is determined by the policy that minimizes the average expected power not supplied over an infinite time horizon. This choice of performance criterion is motivated by two reasons. First, we are interested in strategic, i.e., long-term, decisions and thus consider an infinite time horizon. Second, as power cannot be stored easily (in large quantities), an unsupplied megawatt of power in the future has the same value as an unsupplied megawatt of power today. Thus, the power not supplied (which represents the cost rate in the continuous-time Markov decision process) should not be discounted.

The chapter is organized as follows. The model of the power grid, the considered types of failures, and the (human) resources are described in Section 5.1. In Section 5.2, the continuous-time Markov decision process of the base model, where all states of the system are modeled in detail, is formulated and analyzed. The aggregation of states and the corresponding approximate model is discussed in Section 5.3. Finally, in Section 5.4, both the base model and the aggregated model are applied to a real high-voltage power grid.

5.1 Power grid, failures, and repair resources

We consider a power supply system consisting of a power grid (power supply lines, transformers, busbars), consumers/loads, and the (human) resources of a power grid operator. Via the power grid, the electrical energy is supplied to the consumers. Due to technical failures or external causes, the electrical equipment can fail and thus has to be repaired by the resources.

We would like to remark at this point that the power grid model used in this chapter is simplified in many aspects. State-of-the-art power grid models are far more sophisticated. Our aim is to demonstrate – in a simplified setting – how the effects of a limited availability of resources
on the continuity of supply in redundant power grids can be analyzed and quantified. At the end of Section 5.1, we further discuss some limitations of our power grid model.

### 5.1.1 Power grid model

A schematic view of an extract of a power grid as considered in this model is given in Figure 5.1. The power grid consists of three voltage levels: extra-high-voltage, high-voltage, and medium-voltage. The focus lies on the high-voltage power grid. It is assumed that the power is generated in the extra-high-voltage power grid and fed into the high-voltage power grid via transformers. In the high-voltage power grid, the power is distributed. The medium-voltage power grid represents the consumers. The consumers are aggregated to so-called load points that are connected to the high-voltage power grid via transformers. The model also allows load points being directly connected to the high-voltage power grid.

The power grid is represented by an undirected graph \((V, E)\) with node set \(V\) and edge set \(E\). This is convenient for calculating the (approximate) power flow in the grid with a model based on the DC power flow (cf. Section 5.2.5). The node set \(V\) is given by \(V = V_{\text{trafo}} \cup V_{\text{feed}} \cup V_{\text{load}} \cup V_{\text{busbar}} \cup V_{\text{aux}}\). \(V_{\text{trafo}}\) is the set of all transformers. The transformers connect the different voltage levels by transforming the voltage to the appropriate level. The set of feed-in points \(V_{\text{feed}}\) represents the extra-high-voltage power grid, where the power is generated. It is assumed that the extra-high-voltage power grid is 100% reliable and that each feed-in point can provide a “sufficient” amount of power. The set of load points \(V_{\text{load}}\) represents the consumers and their demand. It is assumed that the demand of each load point is constant over time.\(^1\) \(V_{\text{busbar}}\) contains all busbars. The set of power supply lines is denoted by \(L\). A power supply line \(l \in L\) may have more than two endpoints (cf. Figure 5.1). For the graph representation, such a power supply line is split into several line segments by adding auxiliary nodes. A line segment is a connection between two nodes of \(V\) and the edge set \(E\) of the graph is the set of all line segments. \(V_{\text{aux}}\) contains the aforementioned auxiliary nodes. Each power supply line \(l \in L\) thus corresponds to a subset \(E_l \subseteq E\) of line

\(^1\) This assumption guarantees that the cost in the Markov decision process do not depend on time.
segments (edges). An auxiliary line segment $e \in E$ is a line segment that connects a transformer, a feed-in point, or a load point to a busbar.

As mentioned in Section 2.1.3, the high-voltage power grid is built with redundancies. The redundancies stem from the meshed structure of the grid as well as from the installation of several transformers, busbars, or power supply lines in parallel. In our model, a transformer $v \in V_{\text{trafo}}$, a busbar $v \in V_{\text{busbar}}$, or a power supply line $l \in L$ can therefore represent several identical transformers, busbars, or power supply lines, respectively, that are installed in parallel. We refer to these identical parallel elements as redundant elements of the corresponding node/edge. In Figure 5.1, the numbers in brackets indicate the number of redundant elements that are represented by the corresponding node/edge. If no number is provided, the corresponding node/edge consists of a single element. We assume that the redundancy is active, i.e., all installed elements are connected and used as long as they have not failed (cf. Section 2.1.3). Figure 5.2 illustrates how the active redundancy of redundant elements is modeled. If the parallel elements are not identical (e.g., if two trans-
5.1 Power grid, failures, and repair resources

Figure 5.2  Modeling of redundancies in (i) a transformer, (ii) a busbar, and (iii) a power supply line.

formers that are installed in parallel have different capacities (cf. Section 5.2.5)), they have to be modeled individually. The two transformers at the bottom right of Figure 5.1 are assumed to have different capacities and thus have to be modeled by two separate nodes $v$ and $v'$ in $V_{\text{trafo}}$.

5.1.2 Failures, failing components and resources

We consider two types of failures that can occur in the power grid, namely independent failures and common mode failures. Independent failures affect exactly one transformer, busbar, or power supply line and cause the failure of exactly one of the corresponding redundant elements (independently of other elements of the power grid), e.g., due to a technical failure. Common mode failures can affect one or several power supply lines at the same time and cause the failure of all of the redundant elements of these power supply lines. For instance, if a lightning strike hits a tower that carries two overhead lines, both lines may be affected. A transformer, a busbar, or a power supply line is called operating as long as at least one of its redundant elements is operating (i.e., has not failed). A transformer, a busbar, or a power supply line can be affected by an independent failure as long as it is operating. A common mode failure can occur as long as at least one of the affected power supply lines is operating.

As the extra-high-voltage power grid is assumed to be 100% reliable, feed-in points (and the corresponding auxiliary line segments) do not fail. Also, as the demand of each load point is assumed to be constant, load points (and the corresponding auxiliary line segments) do not fail. Without loss of generality, it can be assumed that the other auxiliary line segments and auxiliary nodes do not fail either, namely by including their
failures in the failure characteristics of the corresponding transformer or power supply line, respectively.

To model the failures in the power grid, we introduce the abstract notion of (failing) components. A (failing) component \( c \) corresponds

(i) to a busbar \( v_c \in V_{\text{busbar}} \), a transformer \( v_c \in V_{\text{trafo}} \), or a power supply line \( l_c \in L \) that can fail independently, or

(ii) to a set of power supply lines \( L_c \subseteq L \) that can fail due to a common mode failure.

Let \( C_{\text{indep}} \) and \( C_{\text{CM}} \) denote the set of failing components of type (i) and (ii), respectively. The set of all failing components is denoted by \( C \), i.e., \( C := C_{\text{indep}} \cup C_{\text{CM}} \). We assume that each component \( c \in C_{\text{indep}} \) corresponds to a distinct busbar, transformer, or power supply line, respectively. Further, define \( C_{\text{indep,line}} := \{ c \in C_{\text{indep}} \mid c \text{ corresponds to some } l_c \in L \} \), i.e., \( C_{\text{indep,line}} \) collects those components of \( C_{\text{indep}} \) that correspond to a power supply line. Examples of failing components are illustrated in Figure 5.3.

For a component \( c \in C_{\text{CM}} \) which corresponds to the set \( L_c \subseteq L \) of power supply lines, define

\[
E_c := \bigcup_{l \in L_c} E_l ,
\]

i.e., \( E_c \) contains the line segments that are affected by the common mode failure corresponding to \( c \). We make the following (natural) assumption:

For each “common mode” component \( c \in C_{\text{CM}} \) and each of the corresponding power supply lines \( l \in L_c \), there exists an “independent” component \( c' \in C_{\text{indep,line}} \) that corresponds to \( l \), i.e., \( l_{c'} = l \).

Assumption (5.1) assures that a power supply line can only be affected by a common mode failure if it is also affected by independent failures. Consequently, a component \( c \in C_{\text{CM}} \) corresponds to the same power supply lines as some set of components \( C_c \subseteq C_{\text{indep,line}} \), i.e.,

\[
E_c = \bigcup_{c' \in C_c} E_{l_{c'}} .
\]

Recall, however, that the effects of the corresponding failures are different. A common mode failure of a power supply line \( l \in L \) leads to
Figure 5.3 Examples of failing components. Components $c_1$ and $c_2$ correspond to the power supply lines $l_1$ (with two redundant elements) and $l_2$, respectively, and component $c_3$ corresponds to busbar $v_1$ (with two redundant elements). $c_1$, $c_2$, and $c_3$ model independent failures of $l_1$, $l_2$, and $v_1$, respectively. Component $c_4$ corresponds to $l_1$ and $l_2$ and models common mode failures affecting these two lines. Thus, in this example, $C^{\text{indep}} = \{c_1, c_2, c_3\}$, $C^{\text{CM}} = \{c_4\}$, and $C^{\text{indep,line}} = \{c_1, c_2\}$.

the failure of all redundant elements of $l$, whereas an independent failure causes the failure of only one of the redundant elements of $l$. For a component $c \in C^{\text{CM}}$, the set $C_c$ is characterized by

$$C_c = \{c' \in C^{\text{indep,line}} | E_{l,c} \cap E_{c'} \neq \emptyset\}. \quad (5.3)$$

If an independent failure or a common mode failure has occurred, the affected elements of the power grid have to be repaired. The repair is performed by a finite number $n_{\text{res}}$ of (human) resources. We neglect travel times of the resources and assume that the resources can instantly move (“jump”) between the components.\(^2\) This assumption is justified as long as the average travel time is significantly smaller than the average repair time, i.e., as long as the considered supply area is not too large. Due to this assumption, the locations and the traveling of the resources do not have to be tracked, which allows to describe the system with fewer states.

It is assumed that the $n_{\text{res}}$ resources are equally qualified and that a single resource can repair

- one redundant element of a component $c \in C^{\text{indep}}$ that has failed due to an independent failure, or
- all redundant elements of a component $c \in C^{\text{CM}}$ that have failed due to the corresponding common mode failure.

\(^2\) In [Zür08, Zür09], two similar models are presented where travel times are included, cf. the comments towards the end of this section.
Thus, a component $c \in C^{\text{indep}}$ where several redundant elements have failed requires several resources independently of each other. On the other hand, a resource can repair several elements of the power grid at once, if they have failed due to the same common mode failure.

**Remark 5.1.1** Our notion of failing components assures that any (independent or common mode) failure and the work of any resource influences exactly one failing component. This property is particularly useful for the aggregated model (cf. (5.24)).

We assume that the following quantities are random variables with an exponential distribution (with a positive parameter; cf. (4.9)):

- the time until the next independent failure of an (operating) redundant element of a component $c \in C^{\text{indep}}$,
- the time needed to repair an independently failed redundant element of a component $c \in C^{\text{indep}}$,
- the time until the next occurrence of a common mode failure corresponding to a component $c \in C^{\text{CM}}$ (when it can occur), and
- the time needed to repair a common mode failure corresponding to a component $c \in C^{\text{CM}}$ that has occurred.

The assumption of exponential distributions is essential for the formulation of the model as a continuous-time Markov decision process (cf. Section 5.2.3). To obtain a more realistic model, different distributions would have to be used. In this case, the model might be formulated as a semi-Markov decision process, cf. Remark 4.2.4 on page 85.

We further assume that all times until failure and all repair times are independent. In particular, the redundant elements of a single transformer, busbar, or power supply line thus fail independently of each other. As the redundant elements of a component $c \in C^{\text{indep}}$ are identical, the exponential distributions of their failure and repair times can all be characterized by the same parameters $\alpha^{\text{fail}}_c > 0$ and $\alpha^{\text{repair}}_c > 0$, respectively. For a component $c \in C^{\text{CM}}$, let $\alpha^{\text{fail}}_c > 0$ and $\alpha^{\text{repair}}_c > 0$ denote the parameters of the exponential distribution of the failure and repair time, respectively, of the corresponding common mode failure. $\alpha^{\text{fail}}_c$ is referred to as failure rate of $c \in C$ and $\alpha^{\text{repair}}_c$ as repair rate of $c \in C$. We further assume that for each $c \in C$, the parameters $\alpha^{\text{fail}}_c$ and $\alpha^{\text{repair}}_c$ are constants. In particular, they do not vary with time. We thus implicitly assume that the power grid is sufficiently maintained such that there is no wearout.
In [Zür08, Zür09], two similar models are presented based on discrete-time Markov decision processes. Both theses consider independent failures only. In [Zür08], the repair of a component takes a deterministic number of time steps, and in [Zür09], the repair is modeled by Bernoulli random variables whose expected values equal the expected repair times. In addition, travel times of the resources are included in both works. [Zür08] uses randomized traveling (with Bernoulli random variables) and [Zür09] assumes a constant travel time of 1 time step between any two components. In contrast to these two approaches, we model the system in continuous-time (by a continuous-time MDP) and, as already mentioned, we neglect travel times to have a description with fewer states.

The power grid model that is described in this section contains many simplifications. In state-of-the-art reliability evaluations, the power grid models are far more detailed and sophisticated. In reality, for instance, not all pieces of equipment are permanently connected to the power grid and these standby components are only switched on if needed. Further we assume a constant demand in each load point. More realistically, the time-dependency of each load point should be included. The possibility of resupplying consumers via emergency power generators is also neglected. We consider only two types of failures (independent failures and common mode failures). State-of-the-art reliability evaluations incorporate many more types of failures, e.g., multiple faults, failures in protection devices, different types of common mode failures and also include cascading effects of failures, etc. Finally, assuming exponential distributions for the failure and repair times is a further simplification.

5.2 Base model

In this section we show how to model the power supply system by a continuous-time Markov decision process

\[ \mathcal{M} = \{ S, U(x), \alpha_{x,u}, p(\cdot \mid x, u), g(x, u) \} \]

as described in Section 4.2.1.\(^3\) We first define the state space \( S \) and the sets of controls \( U(x), x \in S \) (Sections 5.2.1 and 5.2.2). In Section 5.2.3 we specify the rates of leaving \( \alpha_{x,u} \) and the jump probabilities \( p(\cdot \mid x, u) \) and show that the state indeed evolves according to a continuous-time Markov

\(^3\) For notational convenience, we omit the superscript “cont” in \( \mathcal{M} \).
decision process. In Section 5.2.4, we analyze the structural properties of the continuous-time Markov decision process $M$ via its uniformization. Finally, in Section 5.2.5, the calculation of the power not supplied with the DC load shedding model is explained and the expected cost $g(x, u)$ are specified.

5.2.1 State of the system $x$

The state $x_c$ of a failing component $c \in C$ shall represent the number of required resources in this component. As will become clear in the following, this state representation allows to uniquely determine the "state" of the power grid, i.e., how many redundant elements of the power grid are operating.

For a (failing) component $c \in C^{\text{indep}}$, let $m_c \in \mathbb{N}$ denote the total number of corresponding redundant elements (i.e., identical transformers, busbars, or power supply lines that are installed in parallel). A component $c \in C^{\text{indep}}$ requires $k$ resources if and only if $k$ of the $m_c$ redundant elements have failed due to independent failures. Thus, the state $x_c$ of a component $c \in C^{\text{indep}}$ can be represented by

$$x_c \in S_c := \{0, 1, \ldots, m_c\}.$$  

A (failing) component $c \in C^{\text{CM}}$ requires one resource if and only if the corresponding common mode failure has occurred. The state $x_c$ of a component $c \in C^{\text{CM}}$ can thus be represented by

$$x_c \in S_c := \{0, 1\}.$$  

If $x_c = 1$ for $c \in C^{\text{CM}}$, all redundant elements of the corresponding power supply lines have failed. For convenience, we define $m_c := 1$ for $c \in C^{\text{CM}}$.

We enumerate the components such that $C = \{c_1, \ldots, c_{|C|}\}$. The state $x$ of the system is given by

$$x := (x_{c_1}, \ldots, x_{c_{|C|}})$$

and belongs to the state space

$$S := S_{c_1} \times \cdots \times S_{c_{|C|}}.$$
The total number of states is

$$|S| = \prod_{c \in C} (m_c + 1).$$

Hence $|S|$ grows at least as fast as $2^{|C|}$ (because $m_c \geq 1$ for all $c \in C$), i.e., $|S|$ grows exponentially with the number of components.

### 5.2.2 Control $u$

For a given state $x \in S$, a control has to decide how many resources to assign to each component $c \in C$ such that the number of available resources $n_{\text{res}}$ is not exceeded. For a given state $x \in S$, a control can thus be represented by a vector

$$u := (u_1, \ldots, u_{|C|}) \in \mathbb{N}_{0}^{[C]}$$

where $u_c$ indicates how many resources are assigned to component $c \in C$. For $x \in S$, the total number of required resources is given by $\sum_{c \in C} x_c$. An allowable control has to satisfy

$$u_c \in \{0, 1, \ldots, x_c\} \quad \text{for all } c \in C,$$

i.e., no component receives more resources than it requires, and

$$\sum_{c \in C} u_c = \min \{ \sum_{c \in C} x_c, n_{\text{res}} \}. \quad (5.5)$$

The left hand side of the last constraint counts the total number of assigned resources. The constraint makes sure that at most $n_{\text{res}}$ resources are assigned and that no resource remains idle if it could be used for a repair task. For a state $x \in S$, the set of allowable controls is defined by

$$U(x) := \{ u \in \mathbb{N}_{0}^{[C]} \mid u \text{ satisfies } (5.4), (5.5) \}.$$

Let $U := \bigcup_{x \in S} U(x)$. The set of policies with deterministic Markovian decision rules $\Pi^{\text{MD}}$ is defined by

$$\Pi^{\text{MD}} := \{ \pi = \{\mu_0, \mu_1, \ldots\} \mid \mu_n : S \to U, x \mapsto \mu_n(x) \in U(x) \forall x \in S, \forall n \in \mathbb{N}_{0} \}.$$
As will turn out, the continuous-time Markov decision process $\mathcal{M}$ is unichain (for $n_{\text{res}} \geq 1$; cf. Section 5.2.4) and thus it suffices to consider deterministic stationary policies $\mu \in \Pi^{MD}$.\(^4\)

For a state $x \in S$ where enough resources are available (i.e., $n_{\text{res}} \geq \sum_{c \in C} x_c$), there is exactly one control (i.e., $|U(x)| = 1$). In the following, we describe a procedure to calculate the number of controls $|U(x)|$ for a state $x \in S$ where too few resources are available (i.e., $n_{\text{res}} < \sum_{c \in C} x_c$).

We need the following general definition. Let $n \in \mathbb{N}$. For a vector $y \in \mathbb{N}^n_{\leq 0}$, $i \in \mathbb{N}$, $i \leq n$, and $j \in \mathbb{N}_0$, we define

$$N_{ij}(y) := \left| \{ y' \in \mathbb{N}^n_{\leq 0} \mid y'_k \leq y_k \ \forall k, y'_i = 0 \text{ for } 1 \leq k \leq i - 1, \sum_{k=i}^n y'_k = j \} \right|,$$

i.e., $N_{ij}(y)$ is the number of possibilities to distribute $j$ “integer units” to the $n - i + 1$ components $y'_i, \ldots, y'_n$ without exceeding the bounds $y'_k \leq y_k$, $i \leq k \leq n$, given by $y$. For all $j \in \mathbb{N}_0$, $N_{ij}(y)$ can be calculated recursively by

$$N_{nj}(y) = \begin{cases} 1 & \text{if } j \leq y_n \\ 0 & \text{otherwise,} \end{cases} \quad N_{ij}(y) = \sum_{l=0}^{\min\{j,y_i\}} N_{i+1,j-l} \quad \text{for } i = n - 1, \ldots, 1. \quad (5.6)$$

In the sum of (5.6), $l$ is the number of “integer units” that are assigned to component $y'_i$.

For a given state $x \in S$ (i.e., $x \in \mathbb{N}^{|C|}_{\leq 0}$), $N_{ij}(x)$ can be interpreted as the number of possibilities to assign exactly $j$ resources to the components $c_i, \ldots, c_{|C|}$ such that component $c_k$, $i \leq k \leq |C|$, receives at most $x_k$ resources. If too few resources are available to repair all failures in $x$ (i.e., if $n_{\text{res}} < \sum_{c \in C} x_c$), a control $u \in U(x)$ assigns exactly $n_{\text{res}}$ resources and we have

$$|U(x)| = N_{1,n_{\text{res}}}(x).$$

**Remark 5.2.1** If no resource is available, i.e., if $n_{\text{res}} = 0$, there is exactly one control in each state, namely $u^0 := (0, \ldots, 0) \in \mathbb{N}^{|C|}_{\leq 0}$.\(^4\) cf. the comment after Theorem 4.2.17.
5.2.3 State transitions

We first consider the possible transitions of the state $x_c$ of a single component $c \in C$. For a component $c \in C^{\text{indep}}$, the possible transitions of $x_c$ are shown in Figure 5.4(i). As long as the corresponding transformer, busbar, or power supply line is operating, every (additional) independent failure increases the number of failed redundant elements and thus the number of required resources by 1. Every completed repair reduces the number of failed redundant elements and thus the number of required resources by 1. For a component $c \in C^{\text{CM}}$, the possible transitions of $x_c$ are shown in Figure 5.4(ii). If the corresponding common mode failure occurs, one resource is required. This resource is required until the repair of the common mode failure is completed.

\begin{figure}[h]
\centering
\begin{tikzpicture}
\node (start) at (0,0) [state] {\(x_c = 0\)};
\node (1) [state, right of=start] {\(x_c = 1\)};
\node (m-1) [state, right of=1] {\(x_c = m_c - 1\)};
\node (m) [state, right of=m-1] {\(x_c = m_c\)};
\node (end) [state, below of=m, yshift=-1cm] {\(x_c = 0\)};
\node (end2) [state, below of=m, yshift=-1cm] {\(x_c = 1\)};
\path [->]
(start) edge [loop above] node [above] {repair} (start)
(start) edge [bend left] node [above] {independent failure} (1)
(1) edge [bend left] node [above] {repair} (m-1)
(m-1) edge [bend left] node [above] {independent failure} (m)
(m) edge [bend left] node [above] {repair} (end)
(end) edge [loop below] node [below] {repair} (end);
\end{tikzpicture}
\caption{Possible transitions of the state $x_c$ of (i) a component $c \in C^{\text{indep}}$ and (ii) a component $c \in C^{\text{CM}}$.}
\end{figure}

The following example illustrates transitions of the (whole) state of the system.
Example 5.2.2 Consider the power grid and failing components that are depicted in Figure 5.5 (which are identical to those in Figure 5.3). The failing components \( c_1, c_2, \) and \( c_3 \) model independent failures of \( l_1, l_2, \) and \( v_1, \) respectively, and \( c_4 \) models common mode failures affecting \( l_1 \) and \( l_2. \) The power supply line \( l_1 \) and the busbar \( v_1 \) both consist of two redundant elements. We assume that there is one resource available, i.e., \( n_{\text{res}} = 1. \)

![Power grid and failing components](image)

**Figure 5.5** Power grid and failing components. Two independent failures have occurred, one in \( c_2 \) and one in \( c_3, \) and one resource is available for the repair work.

We start from the state where 2 independent failures have occurred, one in \( c_2 \) and one in \( c_3, \) and assign the resource to component \( c_3. \) The possible state transitions are illustrated below. The alternative control of assigning the resource to \( c_2 \) and the alternative transition (instead of “repair of \( c_3 \))” are marked with dashed arrows.

\[
\begin{align*}
\text{independent failure in } c_1 \quad &\rightarrow (1, 1, 1, 0) \\
\text{independent failure of second redundant element in } c_3 \quad &\rightarrow (0, 1, 2, 0) \\
\text{common mode failure in } c_4 \quad &\rightarrow (0, 1, 1, 1) \\
\text{repair of } c_3 \quad &\rightarrow (0, 1, 0, 0) \\
\text{(repair of } c_2) \quad &\rightarrow (0, 0, 1, 0)
\end{align*}
\]
In the sequel, we show that due to our assumption that the times until failures and repair times are exponentially distributed, the state \( x = (x_{c_1}, \ldots, x_{c_{|C|}}) \) of the system indeed evolves according to a continuous-time Markov decision process as described in Section 4.2.1.

We assume we can adapt the control whenever the state of the system changes, i.e., the decision epochs equal the points in time when a transition to a new state occurs. Intuitively, the state evolves as follows. At a given point in time, for each failure (independent or common mode) that can occur and each repair task that is worked on, a (random) timer is set that is distributed like the corresponding time until failure or repair time. The timer that elapses first determines the subsequent state of the system. Due to the memoryless property of exponential distributions,\(^5\) these timers can be reset at any (new) point in time without changing the probabilistic behavior of the system.

For a state \( x \in S \) and a control \( u \in U(x) \), let \( \mathcal{T}(x, u) \) denote the set of all “such timers”, i.e., \( \mathcal{T}(x, u) \) is the collection of random variables of the times until failure and the repair times that correspond to state \( x \) and control \( u \). We call \( \mathcal{T}(x, u) \) the set of active timers (for \( x \) and \( u \)). To specify which random variables belong to \( \mathcal{T}(x, u) \), we need the following definitions. For a state \( x \in S \) and a component \( c \in C_{\text{indep}} \), we define

\[
\chi_{c}(x) := \begin{cases} 
1 & \text{if the corresponding transformer, busbar, or power supply line is operating} \\
0 & \text{otherwise},
\end{cases}
\]

and for a component \( c \in C_{\text{CM}} \), we define

\[
\chi_{c}(x) := \begin{cases} 
1 & \text{if at least one affected power supply line is operating} \\
0 & \text{otherwise}.
\end{cases}
\]

For a component \( c \in C \), \( \chi_{c}(x) \) specifies whether or not an independent failure (if \( c \in C_{\text{indep}} \)) or a common mode failure (if \( c \in C_{\text{CM}} \)) can affect this component.\(^6\) We thus define the set of components that can be

---

\(^5\) See [Nor97, Theorem 2.3.1]: \( T \sim \text{EXP}(\lambda) \) if and only if \( \mathbb{P}[T > s + t \mid T > s] = \mathbb{P}[T > t] \) for all \( s, t \in \mathbb{R}_{\geq 0} \).

\(^6\) For each \( c \in C \), we have \( \chi_{c}(x) = 0 \) whenever \( x_{c} = m_{c} \). For a component \( c \in C_{\text{indep}} \setminus C_{\text{indep, line}} \) (i.e., \( c \) corresponds to a busbar or a transformer), we have \( \chi_{c}(x) = 1 \) if and only if \( x_{c} \neq m_{c} \). For all other components, however, the value of \( \chi_{c}(x) \) cannot always be deduced from \( x_{c} \). For a component \( c \in C_{\text{CM}} \) with \( x_{c} = 0 \), \( \chi_{c}(x) \) can be 0 (if all affected power supply lines have already failed) or 1. Also, if \( x_{c} \neq m_{c} \) for a component \( c \in C_{\text{indep, line}} \), \( \chi_{c}(x) \) can be 0 (if the power supply line has already failed due to a common mode failure) or 1.
affected by failures
\[ C^{\text{fail}}(x) := \{ c \in C \mid \chi_c(x) = 1 \} . \]

For a state \( x \in S \), we further define the set of components where an (independent or common mode) failure has occurred
\[ C^{\text{failed}}(x) := \{ c \in C \mid x_c > 0 \} , \]
and for \( u \in U(x) \), we define the set of components that are assigned at least one resource for repair
\[ C^{\text{repair}}(x, u) := \{ c \in C^{\text{failed}}(x) \mid u_c > 0 \} . \]

For a state \( x \in S \) and a control \( u \in U(x) \), \( T(x, u) \) contains
- timers of independent failures that can occur, i.e., for each \( c \in C^{\text{indep}} \cap C^{\text{fail}}(x) \), \( T(x, u) \) contains the random variables \( T^{\text{fail}}_{c, 1}, \ldots, T^{\text{fail}}_{c, m_c - x_c} \), where \( T^{\text{fail}}_{c, i} \sim \text{EXP}(\alpha^{\text{fail}}_c) \), \( 1 \leq i \leq m_c - x_c \),
- timers of common mode failures that can occur, i.e., for each \( c \in C^{\text{CM}} \cap C^{\text{fail}}(x) \), \( T(x, u) \) contains the random variable \( T^{\text{fail}}_c \sim \text{EXP}(\alpha^{\text{fail}}_c) \),
- timers for the repair of independent failures, i.e., for each \( c \in C^{\text{indep}} \cap C^{\text{repair}}(x, u) \), \( T(x, u) \) contains the random variables \( T^{\text{repair}}_{c, 1}, \ldots, T^{\text{repair}}_{c, u_c} \), where \( T^{\text{repair}}_{c, i} \sim \text{EXP}(\alpha^{\text{repair}}_c) \), \( 1 \leq i \leq u_c \),
- timers for the repair of common mode failures, i.e., for each \( c \in C^{\text{CM}} \cap C^{\text{repair}}(x, u) \), \( T(x, u) \) contains the random variables \( T^{\text{repair}}_c \sim \text{EXP}(\alpha^{\text{repair}}_c) \).

A transition to a new state occurs as soon as the first timer in \( T(x, u) \) elapses. Hence the “holding time” \( H_{x,u} \) in state \( x \in S \) when choosing control \( u \in U(x) \) (i.e., the time until the next state transition/decision epoch) is distributed like \( \min_{T \in T(x, u)} T \). Due to the assumption that all times until failure and all repair times are independent, all timers (random variables) \( T \in T(x, u) \) are independent. Let \( \alpha_T \) denote the parameter of the exponential distribution of the random variable \( T \in T(x, u) \), i.e., \( T \sim \text{EXP}(\alpha_T) \). Using the property that the minimum of two independent exponential random variables is again exponentially

\[ \chi_c(x) = 1, \text{ we have } x_c \neq m_c, \text{ cf. Footnote 6.} \]
distributed,\(^8\) we have

\[ H_{x,u} \sim \text{EXP}(\alpha_{x,u}) \]

where

\[ \alpha_{x,u} := \sum_{T \in \mathcal{T}(x,u)} \alpha_T \]

is the “rate of leaving” state \( x \) when applying control \( u \in U(x) \).

For \( c \in C^\text{fail}(x) \) (or \( c \in C^\text{repair}(x, u) \)), the elapsing of any of the corresponding timers in \( \mathcal{T}(x, u) \) leads to a uniquely defined subsequent state \( x(c) \). The subsequent state \( x(c) \) is unique because the corresponding “failure event” (“repair event”) only affects the state \( x_c \) of component \( c \) and increases (decreases) its current value by 1. Let \( S^\text{subsequent}(x, u) \) denote the set of possible subsequent states from \( x \in S \) when applying control \( u \in U(x) \), i.e.,

\[ S^\text{subsequent}(x, u) := \{ x(c) \mid c \in C^\text{fail}(x) \} \cup \{ x(c) \mid c \in C^\text{repair}(x, u) \}. \]

Observe that \( c \in C^\text{indep} \) can potentially contribute two states to \( S^\text{subsequent}(x, u) \), one for an additional independent failure and one for a completed repair (cf. component \( c_3 \) in Example 5.2.2). If \( n_{\text{res}} \geq 1 \), we have \( \mathcal{T}(x, u) \neq \emptyset \) for all \( x \in S, u \in U(x) \) and thus \( \alpha_{x,u} > 0 \) and \( S^\text{subsequent}(x, u) \neq \emptyset \) for all \( x \in S, u \in U(x) \). On the other hand, if \( n_{\text{res}} = 0 \), there exists a state \( x \in S \) with \( \mathcal{T}(x, u^0) = \emptyset \) (recall \( U(x) = \{ u^0 \} \) for all \( x \in S \), cf. Remark 5.2.1). For instance, in any state \( \hat{x} \in S \) with \( \hat{x}_c = m_c \) for all \( c \in C^\text{indep} \), no further failure can occur and as no resource is available for repair, \( \mathcal{T}(\hat{x}, u^0) = \emptyset \).

Now let \( x \in S, u \in U(x) \) such that \( \mathcal{T}(x, u) \neq \emptyset \). For \( \bar{x} \in S^\text{subsequent}(x, u) \), let \( \mathcal{T}(\bar{x} \mid x, u) \subseteq \mathcal{T}(x, u) \) denote the set of timers whose elapsing leads to \( \bar{x} \). The sets of timers \( \mathcal{T}(\bar{x} \mid x, u), \bar{x} \in S^\text{subsequent}(x, u) \), provide a partition of \( \mathcal{T}(x, u) \). Let \( \bar{x} \in S^\text{subsequent}(x, u) \). The subsequent state of the system is \( \bar{x} \) if and only if one of the timers in \( \mathcal{T}(\bar{x} \mid x, u) \) elapses first. Thus the “jump probability” \( p(\bar{x} \mid x, u) \) is the probability that one of the

\(^8\) Let \( Y_1, Y_2 \) denote two independent\(^{1} \) exponential random variables, i.e., \( Y_1 \sim \text{EXP}(\lambda_1), Y_2 \sim \text{EXP}(\lambda_2) \) for \( \lambda_1, \lambda_2 \in \mathbb{R}_{\geq 0} \). Then, \( \mathbb{P}[\min\{Y_1, Y_2\} > t] = \mathbb{P}[Y_1 > t, Y_2 > t] \overset{(i)}{=} \mathbb{P}[Y_1 > t] \cdot \mathbb{P}[Y_2 > t] \overset{(4.9)}{=} e^{-\lambda_1 t} \cdot e^{-\lambda_2 t} = e^{-(\lambda_1+\lambda_2)t}, \) i.e., \( \min\{Y_1, Y_2\} \sim \text{EXP}(\lambda_1 + \lambda_2) \).
timers in $T(\bar{x} \mid x, u)$ elapses before any other timer, i.e.,

$$p(\bar{x} \mid x, u) = \mathbb{P}\left[ \bigcup_{T \in T(\bar{x} \mid x, u)} \{T < \min_{T' \in T(x, u) \setminus \{T\}} T'\} \right]$$

$$= \mathbb{P}\left[ \min_{T \in T(\bar{x} \mid x, u)} T < \min_{T' \in T(x, u) \setminus T(\bar{x} \mid x, u)} T' \right]$$

$$= \frac{\sum_{T \in T(\bar{x} \mid x, u)} \alpha_T}{\sum_{T' \in T(x, u)} \alpha_{T'}} \mathbb{P}\left[ \min_{T \in T(\bar{x} \mid x, u)} T < \min_{T' \in T(x, u) \setminus T(\bar{x} \mid x, u)} T' \right]$$

The second last equality $(\star)$ follows from another property of exponentially distributed random variables. Let $Y_1, Y_2$ denote two independent exponential random variables, i.e., $Y_1 \sim \text{EXP}(\lambda_1), Y_2 \sim \text{EXP}(\lambda_2)$ for $\lambda_1, \lambda_2 \in \mathbb{R}_{\geq 0}$ with $\lambda_1 + \lambda_2 > 0$. Then, $\mathbb{P}[Y_1 < Y_2] = \frac{\lambda_1}{\lambda_1 + \lambda_2}$.

Clearly, $p(\bar{x} \mid x, u) = 0$ for all $\bar{x} \notin S_{\text{subsequent}}(x, u)$. For $\bar{x} \in S$, we define

$$\alpha(\bar{x} \mid x, u) := \begin{cases} \sum_{T \in T(\bar{x} \mid x, u)} \alpha_T & \text{if } \bar{x} \in S_{\text{subsequent}}(x, u) \\ 0 & \text{otherwise} \end{cases}$$

i.e., the jump probability $p(\bar{x} \mid x, u)$ can be written as

$$p(\bar{x} \mid x, u) = \frac{\alpha(\bar{x} \mid x, u)}{\alpha_{x, u}}.$$ 

$\alpha(\bar{x} \mid x, u)$ is the “transition rate” from $x$ to $\bar{x}$ when applying control $u \in U(x)$ (cf. (4.14) in Section 4.2.1). For $\bar{x} \in S_{\text{subsequent}}(x, u)$, let $c_{\bar{x}} \in C_{\text{fail}}(x)$ or $c_{\bar{x}} \in C_{\text{repair}}(x, u)$ denote the unique component that satisfies $x(c) = \bar{x}$. Then, for $\bar{x} \in S_{\text{subsequent}}(x, u)$, we have

$$\alpha(\bar{x} \mid x, u) = \begin{cases} (m_{c_{\bar{x}}} - x_{c_{\bar{x}}}) \cdot \alpha_{c_{\bar{x}}}^{\text{fail}} & \text{if } c_{\bar{x}} \in C_{\text{fail}}(x) \\ u_{c_{\bar{x}}} \cdot \alpha_{c_{\bar{x}}}^{\text{repair}} & \text{if } c_{\bar{x}} \in C_{\text{repair}}(x, u) \end{cases},$$

and thus

$$p(\bar{x} \mid x, u) > 0 \quad \text{for all } \bar{x} \in S_{\text{subsequent}}(x, u).$$

As a subsequent state $\bar{x}$ differs from the current state $x$ on exactly one component with probability 1, we have $p(x \mid x, u) = 0$ for all $x \in S, u \in$
$U(x)$ with $T(x, u) \neq \emptyset$. If $T(x, u) = \emptyset$, we have $\alpha_{x,u} = 0$ and we define $p(x \mid x, u) := 1$ and $\alpha(\bar{x} \mid x, u) := 0$ for all $\bar{x} \in S$.

Consequently, the state of the system indeed evolves according to a continuous-time Markov decision process as described in Section 4.2.1. The jump probabilities are given by $p(\bar{x} \mid x, u)$, the rates of leaving by $\alpha_{x,u}$, and the transition rates by $\alpha(\bar{x} \mid x, u)$ and $H_{x,u}$ are indeed the holding times. To complete the model, we have to specify the cost based on which the controls are optimized. This is done in Section 5.2.5. Before that, we analyze the structural properties of the continuous-time MDP $M$.

### 5.2.4 Properties of the Markov decision process

The state of the system evolves according to the continuous-time Markov decision process $M$. In this section, we analyze the structural properties of $M$ via its uniformization. For $x \in S$, $u \in U(x)$ with $T(x, u) \neq \emptyset$, we have

$$\alpha_{x,u} = \sum_{T \in T(x,u)} \alpha_T$$

$$= \sum_{\bar{x} \in S_{\text{subsequent}}(x,u)} \sum_{T \in T(\bar{x} \mid x,u)} \alpha_T$$

$$= \sum_{\bar{x} \in S_{\text{subsequent}}(x,u)} \alpha(\bar{x} \mid x, u)$$

$$\leq \sum_{c \in C} (m_c - x_c) \cdot \alpha^\text{fail}_c + \sum_{c \in C} u_c \cdot \alpha^\text{repair}_c$$

$$\leq \sum_{c \in C} m_c \cdot \max\{\alpha^\text{fail}_c, \alpha^\text{repair}_c\} \cdot$$

Clearly, for $x \in S$, $u \in U(x)$ with $T(x, u) = \emptyset$, $0 = \alpha_{x,u} \leq \sum_{c \in C} m_c \cdot \max\{\alpha^\text{fail}_c, \alpha^\text{repair}_c\}$. We thus define (cf. (4.16))

$$\eta_M := f_\eta \cdot \sum_{c \in C} m_c \cdot \max\{\alpha^\text{fail}_c, \alpha^\text{repair}_c\} \quad (5.10)$$
for some (fixed) \( f_\eta \in \mathbb{R}, f_\eta > 1 \) and let \( \tilde{M} \) denote the uniformization of \( M \) with uniformization parameter \( \eta_M \) (cf. (4.21)), i.e.,
\[
\tilde{M} = \{ S, U(x), \tilde{p}(\cdot \mid x, u), \tilde{g}(x, u) \}.
\]
The transition (jump) probabilities \( \tilde{p}(\cdot \mid x, u) \) of the auxiliary discrete-time MDP \( \tilde{M} \) are given, using \( \alpha_{x,u} = \sum_{x' \neq x} \alpha(x' \mid x, u) \), by (cf. (4.17))
\[
\tilde{p}(\bar{x} \mid x, u) = \begin{cases} 
\frac{1}{\eta_M} \cdot \alpha(\bar{x} \mid x, u) & \text{if } \bar{x} \neq x \\
1 - \frac{1}{\eta_M} \cdot \sum_{x' \neq x} \alpha(x' \mid x, u) & \text{if } \bar{x} = x.
\end{cases}
\]
(5.11)
The cost \( \tilde{g}(x, u) \) of \( \tilde{M} \) is specified in (5.23) on page 133. In the following theorem, we prove that the auxiliary discrete-time MDP \( \tilde{M} \) is unichain.

**Theorem 5.2.3** Let \( n_{\mathrm{res}} \geq 1 \). Then, the auxiliary discrete-time MDP \( \tilde{M} \) is unichain, i.e., for every deterministic stationary policy \( \mu \in \Pi_{\mathrm{MD}} \), the embedded Markov chain \( \tilde{X}_\mu \) (of \( \tilde{M} \)) corresponding to \( \mu \) is unichain.

Before we prove Theorem 5.2.3 we discuss its implications. For any choice of the cost \( \tilde{g}(x, u) \), the average expected cost \( J_{\tilde{M},\mu} \) of any deterministic stationary policy \( \mu \in \Pi_{\mathrm{MD}} \) of the auxiliary discrete-time MDP \( \tilde{M} \) is equal for all initial states, i.e., \( J_{\tilde{M},\mu} = \lambda_{\tilde{M},\mu} e \) for some \( \lambda_{\tilde{M},\mu} \in \mathbb{R} \) (cf. Theorem 4.2.7(i)), and the optimal average expected cost \( J^*_{\tilde{M}} \) of \( \tilde{M} \) is equal for all initial states, i.e., \( J^*_{\tilde{M}} = \lambda^*_{\tilde{M}} e \) for some \( \lambda^*_{\tilde{M}} \in \mathbb{R} \) (cf. Theorem 4.2.10). Moreover, for any choice of the cost \( \tilde{g}(x, u) \), the relative value iteration algorithm, applied to the auxiliary discrete-time MDP \( \tilde{M} \), is assured to terminate after a finite number of iterations (cf. (4.18) and Theorem 4.2.12).

As a consequence of Theorem 5.2.3 and (4.36), the (original) continuous-time MDP \( M \) is also unichain (for \( n_{\mathrm{res}} \geq 1 \)). By Theorem 4.2.15, the average expected cost \( J_{M,\mu} \) of any deterministic stationary policy \( \mu \in \Pi_{\mathrm{MD}} \) of \( M \) is equal for all initial states, i.e., \( J_{M,\mu} = \lambda_{M,\mu} e \) for some \( \lambda_{M,\mu} \in \mathbb{R} \). Moreover, by Theorem 4.2.17(iii), we have that
\[
\lambda_{M,\mu} = \lambda_{\tilde{M},\mu}.
\]
(5.12)
In addition (again by Theorem 4.2.17) the optimal average expected cost \( J^*_M \) of \( M \) is equal for all initial states, i.e., \( J^*_M = \lambda^*_M e \) for some \( \lambda^*_M \in \mathbb{R} \), and
\[
\lambda^*_M = \lambda^*_{\tilde{M}}.
\]
(5.13)
Further, the sets of optimal deterministic stationary policies of $\mathcal{M}$ and $\tilde{\mathcal{M}}$ coincide. Summing up, the optimal average expected cost $\lambda^*_M e$ of $\mathcal{M}$ (which we are interested in) is equal to the optimal average expected cost $\lambda^*_{\tilde{\mathcal{M}}} e$ of the auxiliary discrete-time MDP $\tilde{\mathcal{M}}$ (which we compute).

In the sequel, for a deterministic stationary policy $\mu \in \Pi^{MD}$, we refer to $\lambda_{\mathcal{M},\mu}$ and $\lambda_{\tilde{\mathcal{M}},\mu}$ as the average expected cost of $\mu$ of $\mathcal{M}$ and $\tilde{\mathcal{M}}$, respectively. Also, we refer to $\lambda^*_M$ and $\lambda^*_{\tilde{\mathcal{M}}}$ as the optimal average expected cost of $\mathcal{M}$ and $\tilde{\mathcal{M}}$, respectively.

**Remark 5.2.4** If no resource is available, i.e., if $n_{res} = 0$, no failure can be repaired and the system eventually ends up in a state where all redundant elements of each transformer, busbar, or power supply line have failed and remain in this state forever, i.e., this state is absorbing for the corresponding unique (cf. Remark 5.2.1) embedded Markov chain. In general, however, this state is not unique because a power supply line can (completely) fail due to several independent failures or due to common mode failures. Thus, there are generally several states that are absorbing and, as each absorbing state forms a closed recurrent communicating class, $\mathcal{M}$ (and $\tilde{\mathcal{M}}$) is not unichain if $n_{res} = 0$.

In fact, the above Theorem 5.2.3 is a direct consequence of a stronger statement. Namely, not only do all deterministic stationary policies induce a unichain Markov chain $\tilde{X}_\mu$, but – as shown in the following Theorem 5.2.6 – all deterministic stationary policies induce the same (single) recurrent communicating class and the same transient states. As it turns out, the transient states are only an artifact of our modeling of the states. In the definition of the state $x = (x_{c_1}, \ldots, x_{c_{|C|}})$ in Section 5.2.1, we allow all possible combinations of the individual states $x_c$ of the components. However, some of these combinations cannot occur in reality, as we illustrate in the following example.

**Example 5.2.5** Consider again the power grid and failing components of Example 5.2.2 on page 116 (cf. Figure 5.5). The state $(2, 1, 0, 1)$ cannot occur in reality. The common mode failure ($c_4$) can only occur as long as at least one of the power supply lines corresponding to $c_1$ and $c_2$ is operating and $c_1$ and $c_2$ can only fail independently as long as the corresponding power supply lines have not failed due to the common mode failure corresponding to $c_4$.

The above example illustrates the key property of states that cannot occur in reality, namely that there is no “feasible” sequence of common
mode failures that leads to the state in question. A “feasible” sequence of common mode failures has to respect that a common mode failure can only occur if at least one of the affected power supply lines is operating. In particular, the affected power supply lines must not all have failed due to independent failures. The idea of these “feasible” sequences is formalized in property (A) of the following theorem.

For a state $x \in S$ and a component $c \in C^{CM}$, we define (based on the set $C_c \subseteq C_{\text{indep,line}}$ as defined in (5.3) on page 109)

$$C_c(x) := \{ c' \in C_c \mid x_{c'} \neq m_{c'} \}, \quad (5.14)$$

i.e., $C_c(x)$ collects those components in $C_c$ where, in the given state $x$, not all redundant elements have failed due to independent failures. Clearly, $C_c = C_c((0, \ldots, 0))$ for all $c \in C^{CM}$.

**Theorem 5.2.6** Let

$$S^{\text{recurrent}} := \{ x \in S \mid C^{CM} \cap C^{\text{failed}}(x) = \emptyset \}$$

$$\cup \{ x \in S \mid C^{CM} \cap C^{\text{failed}}(x) \neq \emptyset, x \text{ satisfies (A)} \}$$

where:

1. There exists a permutation $\sigma$ of the set $\{1, \ldots, |C|\}$ with

$$\{ c_{\sigma(1)}, \ldots, c_{\sigma(|C^{CM} \cap C^{\text{failed}}(x)|)} \} = C^{CM} \cap C^{\text{failed}}(x)$$

such that (we use the convention that $\bigcup_{i=a}^{b} = \emptyset$ whenever $b < a$)

$$C_{c_{\sigma(j)}}(x) \setminus \bigcup_{k=1}^{j-1} C_{c_{\sigma(k)}}(x) \neq \emptyset \quad 1 \leq j \leq |C^{CM} \cap C^{\text{failed}}(x)|.$$

Let $n_{\text{res}} \geq 1$ and let $\mu \in \Pi^{MD}$ be a deterministic stationary policy. Then, for the embedded Markov chain $\tilde{X}_\mu$ (of $\tilde{M}$) corresponding to $\mu$:

1. $S^{\text{recurrent}}$ is a closed recurrent communicating class.
2. All states in $S^{\text{transient}} := S \setminus S^{\text{recurrent}}$ are transient.

---

10 In the characterization of the states with (A) we explicitly use Assumption (5.1) which guarantees that $C_c((0, \ldots, 0)) = C_c \neq \emptyset$ for all $c \in C^{CM}$.

11 The idea of (A) is that the common mode failures that have occurred in $x$ (corresponding to the components $C^{CM} \cap C^{\text{failed}}(x)$) can be ordered such that this order gives a sequence of failure events that can lead to state $x$ with positive probability.
Remark 5.2.7 Assume $n_{\text{res}} \geq 1$. If $C^{CM} = \emptyset$, i.e., if no common mode failures can occur in the power supply system, then $S^{\text{recurrent}} = S$ in Theorem 5.2.6 and $S$ is a closed recurrent communicating class for any deterministic stationary policy. On the other hand, if $C^{CM} \neq \emptyset$, i.e., if common mode failures can occur in the power supply system, there exists at least one state that is transient for every deterministic stationary policy. To see this, let $c \in C^{CM}$. Any state $x \in S$ with $x_c = 1$ (i.e., the common mode failure has occurred) and $x_{c'} = m_{c'}$ for all components $c' \in C_c((0, \ldots, 0))$ (i.e., all redundant elements that are (potentially) affected by $c$ have failed due to independent failures) violates condition (A) of Theorem 5.2.6 because $C_c(x) = \emptyset$ and thus $x$ is transient.

Remark 5.2.8 For a state $x \in S$, the condition $C^{CM} \cap C^{\text{failed}}(x) = \emptyset$ is easily checked from the values $x_c, c \in C^{CM}$. On the other hand, to check condition (A) of Theorem 5.2.6 for a state $x \in S$, all permutations of the set $C^{CM} \cap C^{\text{failed}}(x)$ have to be considered in principle. However, condition (A) implies the following condition

\[ (B) \quad C_c(x) \neq \emptyset \quad \forall c \in C^{CM} \cap C^{\text{failed}}(x) \]

which is easily checked from the values $x_c, c \in C^{\text{indep}}$. Moreover, if all common mode failures are “disjoint”, i.e., if $C_c \cap C_{c'} = \emptyset$ for all $c, c' \in C^{CM}$, condition (B) is equivalent to (A). We thus define

\[ C^{CM, \text{intersect}} := \{ c \in C^{CM} \mid \exists c' \in C^{CM} \setminus \{ c \} \text{ with } C_c \cap C_{c'} \neq \emptyset \}, \]

and suggest the following procedure to decide whether a state $x \in S$ is recurrent or transient.
if $C_{CM} \cap C_{\text{failed}}(x) = \emptyset$ then
    $x$ is recurrent
else
    if (B) does not hold then
        $x$ is transient
    else
        if $C_{CM,\text{intersect}} = \emptyset$ then
            $x$ is recurrent
        else
            if (A) holds with $C_{CM,\text{intersect}}$ instead of $C_{CM}$ then
                $x$ is recurrent
            else
                $x$ is transient
            end if
        end if
    end if
end if

Due to condition (B) and by considering $C_{CM,\text{intersect}}$, fewer permutations have to be considered than with the original condition (A).

---

Proof of Theorem 5.2.6 (overview): Let $n_{\text{res}} \geq 1$ and let $\mu \in \Pi^{MD}$ be a deterministic stationary policy. The proof consists of the following steps:

1. We prove that each state in $S_{\text{recurrent}}$ is accessible (for the embedded Markov chain $\tilde{X}_\mu$) from any state in $S$ (cf. Figure 5.6).

2. We show that the states $x \in S_{\text{transient}}$ are not accessible from the state $x^0 := (0, \ldots, 0)$ where no failure has occurred (cf. Figure 5.6).

3. As $x^0 \in S_{\text{recurrent}}$ by definition of $S_{\text{recurrent}}$, we conclude from (2) that $S_{\text{recurrent}}$ is closed (for $\tilde{X}_\mu$). To see this, assume there exist states
\[ \hat{x} \in S^{\text{recurrent}} \text{ and } x \in S^{\text{transient}} \text{ such that } \hat{p}(x \mid \hat{x}, \mu(\hat{x})) > 0, \text{ i.e., } \hat{x} \to x \text{ (} x \text{ is accessible from } \hat{x} \text{). Then, as } x^0 \to \hat{x} \text{ by (1), } x^0 \to x \text{ by (A.2), which is a contradiction to (2).} \]

(4) From (1) it follows that all states in \( S^{\text{recurrent}} \) communicate (for \( \tilde{X}_\mu \)). As \( S^{\text{recurrent}} \) is closed (by (3)), \( S^{\text{recurrent}} \) is a communicating class for \( \tilde{X}_\mu \).

(5) By (3) and (4), \( S^{\text{recurrent}} \) is a (finite) closed communicating class. By [Nor97, Theorem 1.5.6], a finite closed communicating class is recurrent and thus part (i) of the theorem is proven.

(6) Let \( x \in S^{\text{transient}} \). As \( x^0 \in S^{\text{recurrent}} \), \( x^0 \) is accessible from \( x \) (by (1)), i.e., starting from \( x \), \( x^0 \) is visited after a finite number of steps with positive probability. However, after visiting \( x^0 \), the chain never returns to \( x \) (by (2)). Consequently, there is a positive probability that starting from \( x \), \( x \) is only visited a finite number of times, i.e., \( x \) cannot be recurrent. Thus, \( x \) is transient and part (ii) of the theorem is proven.

It remains to prove (1) and (2). The proofs are given in Appendix B.2 on page 192. The proof of (1) relies on the fact that no resource is allowed to remain idle if it could be used to repair a failure in the system, cf. (5.5).

\[ \square \]

### Restriction to recurrent states

Let \( n_{\text{res}} \geq 1 \). As \( \tilde{M} \) is unichain (by Theorem 5.2.3), the average expected cost of a deterministic stationary policy does not depend on its transient states (cf. Theorem 4.2.7(iii) and Theorem 4.1.4). Further, all deterministic stationary policies infer the same set of transient states \( S^{\text{transient}} \) (cf. Theorem 5.2.6). We therefore restrict \( \tilde{M} \) to the set of recurrent states \( S^{\text{recurrent}} \) (as defined in Theorem 5.2.6) and define the corresponding discrete-time Markov decision process \( \tilde{M}^{\text{rec}} \) by

\[ \tilde{M}^{\text{rec}} := \{ S^{\text{recurrent}}, U(x), \hat{p}(\cdot \mid x, u), \tilde{g}(x, u) \}. \quad (5.15) \]

Observe that \( \hat{p}(\cdot \mid x, u) \) are indeed probability distributions: as \( S^{\text{recurrent}} \) is closed for any deterministic stationary policy (cf. Theorem 5.2.6), \( \sum_{\hat{x} \in S^{\text{recurrent}}} \hat{p}(\hat{x} \mid x, u) = 1 \) for all \( x \in S^{\text{recurrent}}, u \in U(x) \). The decision rules and policies of \( \tilde{M}^{\text{rec}} \) are the decision rules and policies of \( \tilde{M} \).
(and $\mathcal{M}$) restricted to $S^{\text{recurrent}}$. As a consequence of Theorem 5.2.6, we have:

\begin{equation}
\tilde{\mathcal{M}}^{\text{rec}} \text{ is recurrent.}
\end{equation}

By Theorem 4.2.7(i), the average expected cost $J_{\tilde{\mathcal{M}}^{\text{rec}}, \mu}$ of a deterministic stationary policy $\mu \in \Pi^{\text{MD}}$ (restricted to $S^{\text{recurrent}}$) of $\tilde{\mathcal{M}}^{\text{rec}}$ is equal for all initial states, i.e., $J_{\tilde{\mathcal{M}}^{\text{rec}}, \mu} = \lambda_{\tilde{\mathcal{M}}^{\text{rec}}, \mu} e$ for some $\lambda_{\tilde{\mathcal{M}}^{\text{rec}}, \mu} \in \mathbb{R}$, and we refer to $\lambda_{\tilde{\mathcal{M}}^{\text{rec}}, \mu}$ as average expected cost of $\mu$ of $\tilde{\mathcal{M}}^{\text{rec}}$. Further (by Theorem 4.2.10), the optimal average expected cost $J^*_{\tilde{\mathcal{M}}^{\text{rec}}}$ of $\tilde{\mathcal{M}}^{\text{rec}}$ is equal for all initial states, i.e., $J^*_{\tilde{\mathcal{M}}^{\text{rec}}} = \lambda^*_{\tilde{\mathcal{M}}^{\text{rec}}} e$ for some $\lambda^*_{\tilde{\mathcal{M}}^{\text{rec}}} \in \mathbb{R}$, and we refer to $\lambda^*_{\tilde{\mathcal{M}}^{\text{rec}}}$ as optimal average expected cost of $\tilde{\mathcal{M}}^{\text{rec}}$.

Let $\mu \in \Pi^{\text{MD}}$ denote a deterministic stationary policy. Recall that $\lambda_{\tilde{\mathcal{M}}, \mu}$ denotes the average expected cost of $\mu$ of $\tilde{\mathcal{M}}$. As the average expected cost only depends on recurrent states, we have

\begin{equation}
\lambda_{\tilde{\mathcal{M}}^{\text{rec}}, \mu} = \lambda_{\mathcal{M}, \mu}.
\end{equation}

As a consequence of Theorem 4.2.10, the optimal average expected cost $\lambda^*_{\mathcal{M}}$ and $\lambda^*_{\tilde{\mathcal{M}}^{\text{rec}}}$ of $\mathcal{M}$ and $\tilde{\mathcal{M}}^{\text{rec}}$, respectively, are equal, i.e.,

\begin{equation}
\lambda^*_{\tilde{\mathcal{M}}^{\text{rec}}} = \lambda^*_{\mathcal{M}}
\end{equation}

and the sets of optimal deterministic stationary policies of $\tilde{\mathcal{M}}$ and $\tilde{\mathcal{M}}^{\text{rec}}$ “coincide”, i.e., an optimal deterministic stationary policy of $\tilde{\mathcal{M}}$ restricted to $S^{\text{recurrent}}$ is optimal for $\tilde{\mathcal{M}}^{\text{rec}}$ and an optimal deterministic stationary policy of $\tilde{\mathcal{M}}^{\text{rec}}$, arbitrarily extended to the transient states $S \setminus S^{\text{recurrent}}$, is optimal for $\tilde{\mathcal{M}}$.

Recall that for a deterministic stationary policy $\mu \in \Pi^{\text{MD}}$, $\lambda_{\mathcal{M}, \mu}$ denotes the average expected cost of $\mu$ of $\mathcal{M}$ and $\lambda^*_{\mathcal{M}}$ denotes the optimal average expected cost of $\mathcal{M}$. From (5.12) and (5.13) we conclude

\begin{equation}
\lambda_{\tilde{\mathcal{M}}^{\text{rec}}, \mu} = \lambda_{\mathcal{M}, \mu},
\end{equation}

and

\begin{equation}
\lambda^*_{\tilde{\mathcal{M}}^{\text{rec}}} = \lambda^*_{\mathcal{M}}.
\end{equation}

Furthermore, the sets of optimal deterministic stationary policies of $\mathcal{M}$ and $\tilde{\mathcal{M}}^{\text{rec}}$ “coincide”, i.e., an optimal deterministic stationary policy of $\mathcal{M}$
restricted to $S^{\text{recurrent}}$ is optimal for $\tilde{M}^{\text{rec}}$ and an optimal deterministic stationary policy of $\tilde{M}^{\text{rec}}$, arbitrarily extended to the transient states $S \setminus S^{\text{recurrent}}$, is optimal for $M$.

To calculate the (optimal) average expected cost of $M$ (and $\tilde{M}$), it thus suffices to work with $\tilde{M}^{\text{rec}}$.

### 5.2.5 DC load shedding model for the cost

In this section, we finally specify the expected cost $g(x, u)$ of $M$ and the expected cost rate $\tilde{g}(x, u)$ of $\tilde{M}$ and $\tilde{M}^{\text{rec}}$. For a given state $x \in S$ it might not be possible to satisfy the total demand of the system and some load has to be shed. We assume that the power grid is operated as to minimize the amount of load that is shed, i.e., to minimize the power not supplied. The power not supplied is assessed with the DC load shedding model that is described in the following. A simplified version of this model is also described in [Bal07, BGLZ09, Zür09]. The model is based on the power flow model in [CLDN02].

In reality, if a load exceeds the capacity limit of the corresponding transformer, the protection system disconnects the whole load, i.e., there is no time to partially shed the affected load. However, the protection systems are often configured to react only if the load exceeds e.g. 120% or 130% of the transformer’s capacity. Thus for some time, the load might in fact exceed the 100% value of the transformer’s capacity without being disconnected. During this period, the power grid operator could (in principle) take actions to reduce parts of the load. In our model, we do not allow that the 100% values of the capacities are exceeded but we allow a partial shedding of the loads.

Each load point $v \in V_{\text{load}}$ is characterized by its (constant) demand $p^\text{load}_v \in \mathbb{R}_{>0}$. We assume that two nodes $v, w \in V$, $v \neq w$, are connected by at most one line segment.\(^{12}\) The line segment that connects $v, w \in V$ is denoted by $\{v, w\} = \{w, v\} \in E$. Each line segment $\{v, w\} \in E$ is characterized by its capacity, which limits the amount of power that can flow on this line, and its reactance. The capacity and reactance of a line segment $\{v, w\} \in E$ depend on how many of the (corresponding) identical line segments that are installed in parallel are operating. For a state $x \in S$ and a line segment $\{v, w\} \in E$, let $n_{\{v, w\}}(x) \in \mathbb{N}_0$ denote

\(^{12}\) Recall, however, that a line segment might represent several identical line segments that are installed in parallel.
the number of identical parallel line segments that are operating. Let $c_{\{v,w\}}^{\text{single}} \in \mathbb{R}_{>0}$ denote the capacity and $r_{\{v,w\}}^{\text{single}} \in \mathbb{R}_{>0}$ the reactance$^{13}$ of a single identical line segment of $\{v, w\} \in E$. If $n_{\{v,w\}}(x) > 0$, the capacity of $\{v, w\}$ is given by

$$c_{\{v,w\}}(x) := n_{\{v,w\}}(x) \cdot c_{\{v,w\}}^{\text{single}}$$

and the reactance of $\{v, w\}$ is given by$^{14}$

$$r_{\{v,w\}}(x) := \frac{r_{\{v,w\}}^{\text{single}}}{n_{\{v,w\}}(x)}.$$ 

If all identical parallel line segments have failed (due to several independent failures or a common mode failure), i.e., if $n_{\{v,w\}}(x) = 0$, the line segment $\{v, w\}$ is disconnected from the power grid.

Each transformer $v \in V_{\text{trafo}}$ is characterized by its capacity, i.e., the maximum amount of power it can transform. The capacity of a transformer $v \in V_{\text{trafo}}$ depends on how many of the (corresponding) identical transformers that are installed in parallel are operating. For a state $x \in S$ and a transformer $v \in V_{\text{trafo}}$, let $n_v(x) \in \mathbb{N}_0$ denote the number of identical parallel transformers that are operating. Let $p_v^{\text{single,max}} \in \mathbb{R}_{>0}$ denote the capacity of a single identical transformer of $v$. If $n_v(x) > 0$, the capacity of $v$ is given by

$$p_v^{\text{max}}(x) := n_v(x) \cdot p_v^{\text{single,max}}.$$ 

To incorporate this constraint, the capacity of both adjacent auxiliary lines is set to $p_v^{\text{max}}(x)$ and their reactance is set to $\frac{1}{n_v(x)}$.$^{15}$ If all identical parallel transformers have failed, i.e., if $n_v(x) = 0$, the transformer is disconnected from the power grid. This is realized by disconnecting both adjacent auxiliary line segments from the power grid.

It is assumed that there is no capacity restriction in busbars $v \in V_{\text{busbar}}$ and auxiliary nodes $v \in V_{\text{aux}}$. Hence if a busbar fails independently

$^{13}$ In the electrical engineering literature, the reactance is generally denoted by $X$. We use $r$ to avoid confusion with the state $x$.

$^{14}$ The reason for this is that in DC circuits, to calculate currents and voltages, $n$ identical parallel resistances with resistance $r$ can be replaced by a single component with resistance $\frac{r}{n}$.

$^{15}$ Without loss of generality, it can be assumed that each of the identical transformers is connected by lines with reactance 1.
and there remains at least one identical parallel busbar operating, there is no immediate effect on the power grid. If no identical parallel busbar remains, all (directly) connected line segments are affected. This is modeled by disconnecting them from the power grid. It is important to remark that the line segments are only “temporarily” disconnected for the calculation of the shed load and their disconnection does not affect the state $x$ or the state transitions.

The capacity of an auxiliary line segment that connects a feed-in point or a load point to a busbar is assumed to be unbounded. Without loss of generality, the reactance of these auxiliary line segments is set to 1.

For a given state $x \in S$ of the system, let $E^{\text{off}}(x) \subseteq E$ denote the set of disconnected line segments. The set of operating line segments is $E^{\text{on}}(x) := E \setminus E^{\text{off}}(x)$. The DC load shedding model is based on the DC power flow as described in [WW96, Section 4.1.4], which approximates the power flow in the grid. Two variables are associated with each node $v \in V$: the net real power $p_v \in \mathbb{R}$ and the voltage angle $\theta_v \in \mathbb{R}$. The net real power indicates the net amount of power that flows in or out of a node. The voltage angle can be interpreted as the “voltage potential” of this node. By convention, feed-in points have a non-negative net real power and load points have a non-positive net real power. Transformers, busbars, and auxiliary nodes have a net real power of 0. As the power is preserved within the grid, it has to hold that

\[
\sum_{v \in V} p_v = 0. \tag{5.19}
\]

The values of the net real power $p_v$, $v \in V$, have to satisfy

\[
\begin{align*}
  p_v &\geq 0 & v &\in V_{\text{feed}}, \\
  -p_{v}^{\text{load}} &\leq p_v & v &\in V_{\text{load}}, \\
  p_v &= 0 & v &\in V_{\text{busbar}} \cup V_{\text{trafo}} \cup V_{\text{aux}}.
\end{align*}
\]

We fix an enumeration of $V$ and identify each $v \in V$ with an integer in \{1, 2, \ldots, |V|\}. The vector of net real power $p := (p_1, \ldots, p_{|V|})^T$ and the vector of voltage angles $\theta := (\theta_1, \ldots, \theta_{|V|})^T$ have to satisfy

\[
p = B(x)\theta \tag{5.20}
\]

where $B(x) \in \mathbb{R}^{|V| \times |V|}$ is the admittance matrix of the (configuration of the) power grid corresponding to the state $x$ of the system.\textsuperscript{\textit{16}} $B(x)$ is

\textsuperscript{\textit{16}} cf. [WW96, p. 108].
defined by

$$B(x)_{vw} := \begin{cases} \sum_{k \in V : \{v,k\} \in E^{on}(x)} \frac{1}{r_{\{v,k\}}(x)} & \text{if } v = w \\ \frac{1}{r_{\{v,w\}}(x)} & \text{if } v \neq w \text{ and } \{v,w\} \in E^{on}(x) \\ 0 & \text{if } v \neq w \text{ and } \{v,w\} \notin E^{on}(x). \end{cases}$$

**Remark 5.2.9**

(i) As $(1, \ldots, 1) \cdot B(x) = (0, \ldots, 0)$ for any state $x \in S$, equation (5.19) is satisfied for any $p, \theta$ that satisfy (5.20).

(ii) For a state $x \in S$ where some elements of the power grid have failed, the graph $(V, E^{on}(x))$ might not be connected. In this case, the admittance matrix $B(x)$ has a “block structure” where each block corresponds to a connected subgraph of $(V, E^{on}(x))$ and the set of equations (5.20) can be solved independently for each block/subgraph.

(iii) Let $x \in S$. For notational convenience, we assume that the graph $(V, E^{on}(x))$ is connected. Then, for any $p \in \mathbb{R}^{|V|}$ with $\sum_{v=1}^{|V|} p_v = 0$ and after choosing a reference node $v \in V$ with $\theta_v := \alpha$ for some $\alpha \in \mathbb{R}$, the equation system (5.20) has a unique solution. This statement follows from Theorem A.2.1 (cf. Appendix A.1, page 177) which also provides a characterization of the solution.

For a line segment $\{v, w\} \in E^{on}(x)$, the power flow from $v$ to $w$ is given by

$$\frac{\theta_v - \theta_w}{r_{\{v,w\}}(x)} \quad (5.21)$$

and has to satisfy the line segment’s capacity constraints

$$-c_{\{v,w\}}(x) \leq \frac{\theta_v - \theta_w}{r_{\{v,w\}}(x)} \leq c_{\{v,w\}}(x).$$

The power not supplied is the difference of the total demand and the net real power that is supplied to the loads:

$$\sum_{v \in V_{load}} p_v^{\text{load}} - \sum_{v \in V_{load}} (-p_v) = \sum_{v \in V_{load}} (p_v^{\text{load}} + p_v).$$

---

17 cf. [WW96, p. 107].
18 In fact, the result (appropriately modified) holds for any connected subgraph of $(V, E^{on}(x))$.
19 cf. [WW96, p. 109].
20 As in [CLDN02], we assume that the constraints on the power flow are symmetric. However, this might not be true in reality.
Thus the minimum power not supplied $\kappa(x)$ for a given state $x \in S$ of the system can be assessed by the following linear program

$$
\kappa(x) := \min \sum_{v \in V_{\text{load}}} (p_v^{\text{load}} + p_v) \tag{5.22}
$$

subject to

$$
\begin{align*}
& p = B(x)\theta \\
& -c_{\{v,w\}}(x) \leq \frac{\theta_v - \theta_w}{r_{\{v,w\}}(x)} \leq c_{\{v,w\}}(x) & \{v, w\} \in E^{\text{on}}(x) \\
& p_v \geq 0 & v \in V_{\text{feed}} \\
& -p_v^{\text{load}} \leq p_v \leq 0 & v \in V_{\text{load}} \\
& p_v = 0 & v \in V_{\text{busbar}} \cup V_{\text{trafo}} \cup V_{\text{aux}} \\
& \theta_v \in \mathbb{R} & v \in V .
\end{align*}
$$

Recall that the capacity constraints of the transformers are incorporated as capacity constraints on the corresponding auxiliary line segments.

**Remark 5.2.10** $p = \theta = (0, \ldots, 0)^T$ are always feasible for the linear program (5.22). The corresponding power not supplied is equal to the total demand $\sum_{v \in V_{\text{load}}} p_v^{\text{load}}$ in the system. Also, as $p_v \leq 0$ for $v \in V_{\text{load}}$, the total demand $\sum_{v \in V_{\text{load}}} p_v^{\text{load}}$ is an upper bound on the power not supplied for any feasible solution $p, \theta$. Finally, as $p_v \geq -p_v^{\text{load}}$ for $v \in V_{\text{load}}$, $\kappa(x) \geq 0$ for all $x \in S$.

The minimum power not supplied $\kappa(x)$ is an instantaneous quantity and thus corresponds to a cost rate. The expected cost $g(x, u)$ until the next decision epoch (i.e., the next state transition) given that state $x \in S$ is observed and control $u \in U(x)$ is chosen at the current decision epoch, is given by (cf. (4.22))

$$
g(x, u) := \frac{\kappa(x)}{\alpha_{x,u}} ,
$$

where $\kappa(x)$ is assessed with the linear program (5.22). $g(x, u)$ is the expected amount of energy not supplied until the next decision epoch. The (optimal) average expected cost (cf. (4.23) and (4.24)) thus corresponds to the (optimal) average expected power not supplied.

The discrete-time cost $\tilde{g}(x, u)$ of $\tilde{M}^{\text{rec}}$ (or $\tilde{M}$) is given by (cf. (4.20))

$$
\tilde{g}(x, u) = \kappa(x) \quad x \in S, u \in U(x) . \tag{5.23}
$$
Thus, the discrete-time cost only depend on the current state \( x \) (and are independent of the chosen control \( u \)).

**Remark 5.2.11** If \( n_{res} = 0 \), the system eventually reaches an absorbing state where all redundant elements have failed (cf. Remark 5.2.4). Let \( S^{\text{absorb}} \subseteq S \) denote the set of these absorbing states. As all the states \( x \in S^{\text{absorb}} \) infer the same set \( E^{\text{on}}(x) \) of operating line segments, the power not supplied is equal for all these states. In particular, the state \( \hat{x} \) with \( \hat{x}_c := m_c \) for \( c \in C^{\text{indep}} \) and \( \hat{x}_c := 0 \) for \( c \in C^{\text{CM}} \) belongs to \( S^{\text{absorb}} \) and thus \( \kappa(x) = \kappa(\hat{x}) \) for all \( x \in S^{\text{absorb}} \). The optimal average expected cost of \( \mathcal{M} \) (and \( \tilde{\mathcal{M}}, \mathcal{M}^{\text{rec}} \)) is thus given by \( \kappa(\hat{x}) \), i.e.,

\[
\lambda^*_{\mathcal{M}} = \lambda^*_{\tilde{\mathcal{M}}} = \lambda^*_{\mathcal{M}^{\text{rec}}} = \kappa(\hat{x}).
\]

**Remark 5.2.12** It seems natural to expect that the power not supplied does not decrease when additional failures occur and does not increase when a repair work is finished. In other words, the power not supplied \( \kappa(\cdot) \) should be monotone, i.e., \( \kappa(x) \leq \kappa(x') \) for all \( x, x' \in S \) with \( x \leq x' \) (i.e., \( x_c \leq x'_c \) for all \( c \in C \)). However, as the following example shows, this property does not hold in general.

Consider the following (simplified) power grid with 1 feed-in point (F), 1 load point (L) with a demand of 3 and 3 busbars (\( v_1, v_2, v_3 \)). The busbars are connected by three power supply lines, whose reactances are assumed to be 1. The feed-in point and the load point are connected to \( v_1 \) and \( v_3 \), respectively, with an auxiliary line segment (a).

\[ \begin{align*}
\text{capacity} & \quad c_{\{v_1,v_3\}}(x) = 1 \\
\text{demand} & \quad = 3
\end{align*} \]

In the displayed situation, the demand of the load point (L) cannot be satisfied. Namely, the capacity constraint on the power supply line \( \{v_1, v_3\} \)
requires that $|\theta_{v_1} - \theta_{v_3}| \leq 1$, i.e., at most 1 unit of power can flow on \{v_1, v_3\}. Setting $\theta_{v_3} := 0$ and using that $p_{v_2} = 0$ ($v_2$ is a busbar), we obtain from (5.20) that $\theta_{v_2} = \frac{1}{2}\theta_{v_1}$. As $\theta_{v_3} = 0$, $|\theta_{v_1}| \leq 1$ and thus at most 0.5 units of power can flow on \{v_1, v_2\} and on \{v_2, v_3\}. Hence in total, at most 1.5 units of power can flow from $v_1$ to $v_3$ (and thus from $F$ to $L$). If the line \{v_1, v_3\} is removed, however, the demand of $L$ can be satisfied. For instance, by setting $\theta_{v_3} := 0$, $\theta_{v_2} := 3$, and $\theta_{v_1} := 6$, the capacity constraints on the power supply lines \{v_1, v_2\} and \{v_2, v_3\} are satisfied and 3 units of power flow from $v_1$ to $v_3$ (and thus from $F$ to $L$).

In Appendix A.2 we discuss a (somewhat artificial) condition that assures that $\kappa(\cdot)$ is monotone. To exclude unrealistic situations where the minimum power not supplied decreases when additional failures occur, binary variables could be added to the linear program (5.22) that allow to (manually) switch off these “limiting” elements of the power grid. By additionally minimizing over these binary variables, the minimum power not supplied would be assessed such that the power not supplied is guaranteed to be monotone.

The DC load shedding model described in this section is a static model and does not include “cascading” effects, i.e., the propagation of failures due to e.g. an overload on other elements of the power grid. A DC power flow model including cascading effects is described e.g. in [CLDN02]. A similar model is used in [Bal07, BGLZ09] where repair strategies are developed and analyzed that aim at delaying cascading effects.

5.3 Aggregated model

In this section we construct an aggregated model that allows to calculate an upper bound on the optimal average expected cost of the base model of Section 5.2. The idea of the aggregated model is to aggregate states that require more than (a “threshold” number of) $K$ resources. The motivation behind this is that during most of the time, there are only few failures in the system, i.e., only few resources are required. Rarely, there are many failures simultaneously. However, states with many failures should not be neglected because these states can occur (although with a very small probability) and they may lead to a large interruption of supply. The main advantage of the approach in this section is that it models states that are likely to occur (i.e., that require only few resources) in detail but
avoids the large number of states of the (completely) detailed base model by aggregating unlikely states (i.e., states that require many resources). A similar approach is suggested in [MdSeSG89] in a decision-free setting. In traditional reliability evaluation models, the issue of the large number of states (i.e., the large number of combinations of failures that are possible in principle) is for instance solved by considering combinations of failures only up to a given order\(^{21}\) or with a probability of occurrence of at least a given value.

Recall that for \(n_{\text{res}} = 0\), the optimal average expected cost is easily determined exactly (cf. Remark 5.2.11). Throughout this section we thus assume \(n_{\text{res}} \geq 1\). As argued at the end of Section 5.2.4, it suffices to work with \(\mathcal{M}_{\text{rec}}\) (instead of \(\mathcal{M}\)) to calculate the (optimal) average expected cost of \(\mathcal{M}\). Throughout this section, we thus work with \(\mathcal{S}_{\text{recurrent}}\) and the recurrent discrete-time MDP \(\mathcal{M}_{\text{rec}}\) (cf. (5.15)). Let \(N\) denote the maximum number of required resources due to independent failures in the components, i.e.,

\[
N := \sum_{c \in C_{\text{indep}}} m_c.
\]

Observe that \(N + |C^{CM}|\) is the maximum number of required resources in any state \(x \in S\). The set of (recurrent) states \(\mathcal{S}_{\text{recurrent}}\) of the system can be partitioned into sets with equally many required resources, i.e., \(\mathcal{S}_{\text{recurrent}} = \bigcup_{k=0}^{N+|C^{CM}|} \mathcal{S}^k\) where

\[
\mathcal{S}^k := \{x \in \mathcal{S}_{\text{recurrent}} \mid \sum_{c \in C} x_c = k\}, \quad 0 \leq k \leq N + |C^{CM}|,
\]

contains all states in \(\mathcal{S}_{\text{recurrent}}\) that require \(k\) resources.\(^{22}\) The following theorem shows that only for \(0 \leq k \leq N\), the sets \(\mathcal{S}^k\) are non-empty. The proof is given in Appendix B.2 on page 198.

**Theorem 5.3.1**

(i) For \(0 \leq k \leq N\), \(\mathcal{S}^k \neq \emptyset\).

(ii) For \(k > N\), \(\mathcal{S}^k = \emptyset\).

We note that \(\mathcal{S}^0\) contains exactly one state, namely \(x^0 := (0, \ldots, 0)\), the state where no failure has occurred. The partition of the set of states \(\mathcal{S}_{\text{recurrent}}\) into \(\mathcal{S}^k\), \(0 \leq k \leq N\), is illustrated in Figure 5.7.

\(^{21}\) cf. the second approximate (minimal) cut set method in [BA92, Section 5.3.3].

\(^{22}\) Very similarly, [MdSeSG89] partition the state space according to the number of failures that have occurred in each state.
In the base model, with each additional (independent or common mode) failure, exactly one component requires one resource more and with each repair, exactly one component requires one resource less. Consequently:

From a state \( x \in S^k \), transitions are only possible to states in \( S^{k+1} \) or \( S^{k-1} \) (or to \( x \in S^0 \) due to the uniformization). \((5.24)\)

We define additional states \( s^1, \ldots, s^N \) where \( s^k \), \( 1 \leq k \leq N \), represents all states in \( S^k \), i.e., \( s^k \) is the aggregation of \( S^k \). We further define \( s^0 := x^0 \). As \( S^0 = \{x^0\}, s^0 \) can be interpreted as aggregation of \( S^0 \).

From the point of view of an aggregated state \( s^k \), it is not known which components require a resource and a control can no longer decide which components receive a resource. Thus there is no choice of control in a state \( s^k \) and the (unique) control simply makes sure that the correct number of resources (i.e., \( \min\{k, n_{res}\} \)) is working.

Let \( K \in \mathbb{N}_0, K \leq N \), denote the aggregation threshold, i.e., all states that require at most \( K \) resources are modeled according to the base model (cf. Section 5.2) whereas states that require more than \( K \) resources are aggregated into the states \( s^k, k > K \). The aggregated model is defined by the discrete-time Markov decision process \( \tilde{\mathcal{M}}_K \) (that depends on the parameter \( K \)), given by

\[
\tilde{\mathcal{M}}_K := \left\{ \bigcup_{k=0}^K S^k \cup \{s^1, \ldots, s^N\}, U(x), \tilde{\rho}(\cdot \mid x, u), \tilde{\rho}(\cdot \mid s^k), \tilde{g}(x, u), \tilde{g}(s^k) \right\},
\]

\((5.25)\)

where \( \tilde{\rho}(\cdot \mid s^k) \) and \( \tilde{g}(s^k) \) denote the transition probabilities and the cost, respectively, in aggregated states \( s^k \) and remain to be specified. The decision rules and policies of \( \tilde{\mathcal{M}}_K \) are the decision rules and policies of \( \mathcal{M} \) (and \( \mathcal{M} \)) restricted to \( \bigcup_{k=0}^K S^k \). The cost \( \tilde{g}(x, u) \) are defined as in \((5.23)\). As long as the system is in a “detailed” state \( x \in S^k, 0 \leq k \leq N \), it transitions only to states \( x \in S^{k+1} \) or \( x \in S^{k-1} \).
$K$, the transition probabilities $\tilde{p}(\cdot \mid x, u)$ as defined in (5.11) are used. The transition probabilities to the aggregated state $s^{K+1}$ are defined (for $K < N$) by

$$\tilde{p}(s^{K+1} \mid x, u) := \sum_{\bar{x} \in S^{K+1}} \tilde{p}(\bar{x} \mid x, u) \quad x \in S, u \in U(x). \quad (5.26)$$

Due to (5.24), $\tilde{p}(s^{K+1} \mid x, u) \neq 0$ only for $x \in S^K$. Whenever the system jumps to the aggregated state $s^{K+1}$, it remains in aggregated states until it reaches the aggregated state $s^0$. Upon reaching $s^0 = x^0$, the system switches back to the “detailed” model.

We construct the transition probabilities $\tilde{p}(\cdot \mid s^k)$ such that property (5.24) is preserved, i.e., transitions are only allowed from $s^k$ to $s^{k-1}$, $s^k$, and $s^{k+1}$ and thus (for $1 \leq k \leq N - 1$)

$$\tilde{p}(s^l \mid s^k) := 0 \quad |l - k| > 1. \quad (5.27)$$

The partition of the set of states $S^{\text{recurrent}}$ into the sets $S^k$, $0 \leq k \leq K$, the aggregated states $s^1, \ldots, s^N$, and the transitions between detailed and aggregated states are illustrated in Figure 5.8.

**Remark 5.3.2** For $K = N$, the discrete-time MDP $\tilde{M}_N$ coincides with $\tilde{M}^{\text{rec}}$, i.e., there are no aggregated states in $\tilde{M}_N$.

We also require

$$\tilde{p}(s^{k+1} \mid s^k) > 0 \quad 1 \leq k \leq N - 1, \quad (5.28)$$

$$\tilde{p}(s^{k-1} \mid s^k) > 0 \quad 1 \leq k \leq N. \quad (5.29)$$

These conditions assure that $\tilde{M}_K$ is a recurrent discrete-time MDP, as shown in the following theorem. The proof is provided in Appendix B.2 on page 199.

**Theorem 5.3.3** Let $n_{\text{res}} \geq 1$ and assume (5.28), (5.29) hold. Then, the discrete-time MDP $\tilde{M}_K$, $0 \leq K \leq N$, is recurrent, i.e., for every deterministic stationary policy $\mu \in \Pi^{\text{MD}}$ restricted to $\bigcup_{k=0}^K S^k$, the embedded Markov chain $\tilde{X}_{K,\mu}$ (of $\tilde{M}_K$) corresponding to $\mu$ is irreducible and all states are recurrent.
As a consequence of Theorem 4.2.7(i), the average expected cost $J_{\tilde{M}_K, \mu}$ of a deterministic stationary policy $\mu \in \Pi^{MD}$ (restricted to $\bigcup_{k=0}^{K} S^k$) of $\tilde{M}_K$ is equal for all initial states, i.e., $J_{\tilde{M}_K, \mu} = \lambda_{\tilde{M}_K, \mu} e$ for some $\lambda_{\tilde{M}_K, \mu} \in \mathbb{R}$, and we refer to $\lambda_{\tilde{M}_K, \mu}$ as average expected cost of $\mu$ of $\tilde{M}_K$. In addition (by Theorem 4.2.10), the optimal average expected cost $J^*_{\tilde{M}_K}$ of $\tilde{M}_K$ is equal for all initial states, i.e., $J^*_{\tilde{M}_K} = \lambda^*_{\tilde{M}_K} e$ for some $\lambda^*_{\tilde{M}_K} \in \mathbb{R}$, and we refer to $\lambda^*_{\tilde{M}_K}$ as optimal average expected cost of $\tilde{M}_K$.

The next theorem provides conditions on the transition probabilities $\tilde{p}(s^l \mid s^k)$ and the cost $\tilde{g}(s^k)$ in aggregated states which assure that the optimal average expected cost of $\tilde{M}_K$ provides an upper bound on the optimal average expected cost of $\tilde{M}^{rec}$ and thus also of the base model $\tilde{M}$.

As already mentioned, [MdSeSG89] describe a similar aggregation technique in a decision-free setting. They formulate a class of approximation models (with different “approximation levels”) that provide upper and lower bounds on the average expected cost (of a Markov chain). The approximation model that corresponds to our approach is also discussed in [LM94]. We use some of the ideas of [MdSeSG89] in the proof of Theorem 5.3.4. Their approach mainly relies on finding a lower bound on the stationary probabilities of the states that are modeled in detail. Our proof is further based on a “monotonicity” property of the stationary probability in each “slice” $S^k \cup s^k$ (cf. Lemma B.2.13 on page 217). In contrast to their approach, we do not necessarily assign the maximum cost to all aggregated states to obtain an upper bound. Some partial results of our proof are also provided in [LM94].

Note that $U(x^0) = \{u^0\}$ with $u^0 = (0, \ldots, 0)$. Thus (as $S^0 = \{x^0\}$) there
is no choice of control in $S^0$. In particular, 
\[ \mu(x^0) = u^0 \] (5.30)
for any deterministic stationary policy $\mu \in \Pi^{MD}$.

**Theorem 5.3.4 (Upper bound)** Let $n_{\text{res}} \geq 1$ and $0 \leq K \leq N$. Assume that the transition probabilities $\tilde{p}(s^l \mid s^k)$ in the aggregated states satisfy (5.27), (5.28), (5.29), and

\[
\begin{align*}
\tilde{p}(s^{k+1} \mid s^k) & \geq \max_{x \in S^k, u \in U(x)} \sum_{\tilde{x} \in S^{k+1}} \tilde{p}(\tilde{x} \mid x, u) \quad 1 \leq k \leq N - 1 \\
\tilde{p}(s^{k-1} \mid s^k) & \leq \min_{x \in S^k, u \in U(x)} \sum_{\tilde{x} \in S^{k-1}} \tilde{p}(\tilde{x} \mid x, u) \quad 1 \leq k \leq N.
\end{align*}
\]

Further assume that the cost of the aggregated states satisfy

\[
\begin{align*}
\tilde{g}(s^k) & \geq \max_{x \in S^k, u \in U(x)} \tilde{g}(x, u) \quad 1 \leq k \leq N, \tag{5.31}
\end{align*}
\]

and

\[
\begin{align*}
\tilde{g}(s^{k+1}) & \geq \tilde{g}(s^k) \geq \tilde{g}(x^0, u^0) \quad 1 \leq k \leq N - 1. \tag{5.32}
\end{align*}
\]

Let $\mu \in \Pi^{MD}$ be a deterministic stationary policy. Then the average expected cost $\lambda_{\tilde{M}_K, \mu}$ of $\mu$ of $\tilde{M}_K$ gives an upper bound on the average expected cost $\lambda_{\tilde{M}_{\text{rec}}, \mu}$ (and $\lambda_{M, \mu}$) of $\mu$ of $\tilde{M}_{\text{rec}}$ (and $M$),\(^{23}\) i.e.,

\[
\lambda_{M, \mu} = (5.17) \lambda_{\tilde{M}_{\text{rec}}, \mu} \leq \lambda_{\tilde{M}_K, \mu}. \tag{5.33}
\]

In addition, the optimal average expected cost $\lambda^*_{\tilde{M}_K}$ of $\tilde{M}_K$ gives an upper bound on the optimal average expected cost of $\tilde{M}_{\text{rec}}$ (and thus also of the base model $M$), i.e.,

\[
\lambda^*_{M} = (5.18) \lambda^*_{\tilde{M}_{\text{rec}}} \leq \lambda^*_{\tilde{M}_K}.
\]

\(^{23}\) For $\tilde{M}_K$, $\mu$ is restricted to $\bigcup_{k=0}^{K} S^k$ and for $\tilde{M}_{\text{rec}}$, $\mu$ is restricted to $S^\text{recurrent}$.
Proof of Theorem 5.3.4 (idea): It is sufficient to show the inequality in (5.33) for any deterministic stationary policy \( \mu \in \Pi^{\text{MD}} \). Namely, let \( \mu^*_{\tilde{M}_K} \in \Pi^{\text{MD}} \) denote a deterministic stationary policy whose restriction to \( \bigcup_{k=0}^{K} S^k \) is optimal for \( \tilde{M}_K \) (which exists by Theorem 4.2.10), i.e.,

\[
\lambda_{\tilde{M}_K, \mu^*_{\tilde{M}_K}} = \min_{\mu \in \Pi^{\text{MD}}} \lambda_{\tilde{M}_K, \mu} = \lambda^*_{\tilde{M}_K}.
\]

Then we conclude

\[
\lambda^*_{\hat{M}_{\text{rec}}} = \min_{\mu \in \Pi^{\text{MD}}} \lambda_{\hat{M}_{\text{rec}}, \mu} \leq \lambda^*_{\hat{M}_{\text{rec}}, \mu^*_{\tilde{M}_K}} \leq \lambda_{\tilde{M}_K, \mu^*_{\tilde{M}_K}} = \lambda^*_{\tilde{M}_K}.
\]

For \( K = N \), \( \hat{M}_N \) coincides with \( \hat{M}_{\text{rec}} \) (cf. Remark 5.3.2) and the theorem is trivially true (with equality) in this case. We thus assume in the following that \( K \leq N - 1 \).

In the proof, we distinguish the case \( K = 0 \) from the case \( K > 0 \), as the former case is considerably easier (cf. also Remark 5.3.5). However, in both cases, we follow the same ideas, which we summarize in the following. Let \( \mu \in \Pi^{\text{MD}} \) be a deterministic stationary policy. For \( \hat{M}_K \), \( \mu \) is restricted to \( \bigcup_{k=0}^{K} S^k \) and for \( \hat{M}_{\text{rec}} \), \( \mu \) is restricted to \( S^{\text{recurrent}} \). The idea of the proof of (5.33) is:

(a) As \( \hat{M}_{\text{rec}} \) and \( \hat{M}_K \) both are recurrent discrete-time MDPs (cf. (5.16) and Theorem 5.3.3), the embedded Markov chains \( \hat{X}_\mu \) and \( \hat{X}_{K, \mu} \) of \( \hat{M}_{\text{rec}} \) and \( \hat{M}_K \), respectively, have unique (positive) stationary distributions \( \nu_\mu \) and \( \nu_{K, \mu} \), respectively (cf. Theorem 4.1.4). The average expected cost \( \lambda_{\hat{M}_{\text{rec}}, \mu} \) and \( \lambda_{\hat{M}_K, \mu} \) of \( \mu \) for \( \hat{M}_{\text{rec}} \) and \( \hat{M}_K \), respectively, are given by (cf. Theorem 4.2.7(iii))

\[
\lambda_{\hat{M}_{\text{rec}}, \mu} = \sum_{x \in S^{\text{recurrent}}} \nu_\mu(x) \bar{g}(x, \mu(x)) \tag{5.34}
\]

\[
\lambda_{\hat{M}_K, \mu} = \sum_{k=0}^{K} \sum_{x \in S^k} \nu_{K, \mu}(x) \bar{g}(x, \mu(x)) + \sum_{k=1}^{N} \nu_{K, \mu}(s^k) \bar{g}(s^k). \tag{5.35}
\]

(b) “Exact aggregation”. The transition probabilities in the aggregated states \( \left( \hat{p}^\text{exact}_\mu(s^l | s^k) \right) \) are specified such that the total stationary
probability in each “slice” $S^k \cup s^k$ equals the total stationary probability of $S^k$ in $\tilde{M}^\text{rec}$. We now formalize this idea. Let $\nu^\text{exact}_\mu$ denote the stationary distribution of the embedded Markov chain of $\tilde{M}_K$ corresponding to $\mu$ and with transition probabilities $\tilde{p}^\text{exact}_\mu (s^l | s^k)$. The total stationary probability of a “slice” $S^k \cup s^k$ is defined by

$$\nu^\text{exact}_\mu (x) = \begin{cases} \nu^\text{exact}_\mu (x^0) & k = 0 \\ \nu^\text{exact}_\mu (s^k) + \sum_{x \in S^k} \nu^\text{exact}_\mu (x) & 1 \leq k \leq K \\ \nu^\text{exact}_\mu (s^k) & K + 1 \leq k \leq N \end{cases}$$

(5.36)

and it is shown that for $0 \leq k \leq N$

$$\sum_{x \in S^k} \nu_\mu (x) = \nu^\text{exact}_\mu (k).$$

(5.37)

(c) It is shown that the average expected cost $\lambda^\text{exact}^{\tilde{M}_K, \mu}$ of the “exact aggregation” (i.e., of $\tilde{M}_K$ corresponding to $\mu$ and with transition probabilities $\tilde{p}^\text{exact}_\mu (s^l | s^k)$) provides an upper bound on $\lambda^{\tilde{M}^\text{rec}, \mu}$, i.e.,

$$\lambda^{\tilde{M}^\text{rec}, \mu} \leq \lambda^\text{exact}^{\tilde{M}_K, \mu}.$$ 

(d) However, the transition probabilities $\tilde{p}^\text{exact}_\mu (s^l | s^k)$ cannot be easily obtained (as they basically rely on knowing the stationary distribution $\nu_\mu$ of $\tilde{X}_\mu$ of the base model which we want to avoid).

(e) “Modification”. As (b) and (c) “link” the base model and the aggregated model, we can henceforth focus on properties of the aggregated model. It is proven that the average expected cost $\lambda^{\tilde{M}_K, \mu}$ of $\tilde{M}_K$ corresponding to $\mu$ and with transition probabilities $\tilde{p}(s^l | s^k)$ as specified in the theorem (which can easily be obtained, cf. Theorem 5.3.6 below), provide an upper bound on the average expected cost of the “exact aggregation”, i.e.,

$$\lambda^\text{exact}^{\tilde{M}_K, \mu} \leq \lambda^{\tilde{M}_K, \mu}.$$ 

(f) By combining the result of (c) and (e), we obtain the inequality in (5.33), which finishes the proof.

The proof of (b), (c), and (e) is given in Appendix B.2 on page 200 for $K = 0$ and on page 205 for $K > 0$ (or, more precisely, for $1 \leq K \leq N - 1$,}
5.3 Aggregated model

\[ N \geq 2 \). The proof of the case \( K = 0 \) is based on [Zen10]. Some ideas in the proof of the case \( K > 0 \) are taken from [MdSeSG89] and some partial results are also proven in [LM94]. □

**Remark 5.3.5** Recall that \( S^0 = \{ x^0 \} \). In the case \( K = 0 \) (using the identification \( s^0 = x^0 \)), all states are aggregated and the state space of the aggregated model \( \mathcal{M}_0 \) is given by \( \{ s^0, s^1, \ldots, s^N \} \). There is no choice of control (cf. (5.30)) and all deterministic stationary policies induce the same embedded Markov chain. Further, the stationary distribution of this embedded Markov chain can be computed analytically, cf. Lemma B.2.5 on page 203. Consequently, the optimal average expected cost \( \lambda^*_{\mathcal{M}_0} \) can be computed via (5.35).

By defining (cf. (5.23))

\[
\tilde{g}(s^k) := \max_{x \in \bigcup_{l=0}^k S^l} \kappa(x) \quad 1 \leq k \leq N, \tag{5.38}
\]

the conditions (5.31) and (5.32) of Theorem 5.3.4 on the cost in the aggregated states are satisfied. The following theorem shows how to construct the transition probabilities \( \tilde{p}(s^l \mid s^k) \) in the aggregated states such that the corresponding assumptions of Theorem 5.3.4 are satisfied. We define

\[
N^\text{max} := \sum_{c \in \mathcal{C}} m_c = N + |\mathcal{C}^\text{CM}|.
\]

**Theorem 5.3.6** Let \( n_{\text{res}} \geq 1 \). Let \( \sigma, \hat{\sigma} \) denote permutations of the set \( \{1, \ldots, |\mathcal{C}|\} \) that satisfy

\[
\alpha^{\text{fail}}_{c_\sigma(1)} \leq \ldots \leq \alpha^{\text{fail}}_{c_\sigma(|\mathcal{C}|)},
\]

and

\[
\alpha^{\text{repair}}_{c_\sigma(1)} \leq \ldots \leq \alpha^{\text{repair}}_{c_\sigma(|\mathcal{C}|)},
\]

respectively. For \( 0 \leq k \leq N^\text{max} - 1 \), we define (using the convention that
\[ \sum_{i=a}^{b} \ldots = 0 \text{ whenever } b < a \] \[ \bar{l}(k) := \max\{ \bar{l} \mid \sum_{l=1}^{\mid C \mid} m_{c_\sigma(l)} \geq N^{\text{max}} - k \} \] \[ \bar{m}(k) := N^{\text{max}} - k - \sum_{l=\bar{l}(k)+1}^{\mid C \mid} m_{c_\sigma(l)} , \]

and for \( 1 \leq k \leq N^{\text{max}} \), we define

\[ l(k) := \min\{ \bar{l} \mid \sum_{l=1}^{\bar{l}} m_{c_\sigma(l)} \geq \min\{ k, n_{\text{res}} \} \} \] \[ m(k) := \min\{ k, n_{\text{res}} \} - \sum_{l=1}^{l(k)-1} m_{c_\sigma(l)} . \]

Define the transition probabilities \( \tilde{p}(s^l \mid s^k) \) in aggregated states as follows:

\[ \tilde{p}(s^{k+1} \mid s^k) := \frac{1}{\eta M} \cdot q^{UB,+}_k \quad 1 \leq k \leq N - 1 \] \[ \tilde{p}(s^{k-1} \mid s^k) := \frac{1}{\eta M} \cdot q^{UB,-}_k \quad 1 \leq k \leq N \] \[ \tilde{p}(s^l \mid s^k) := 0 \quad 1 \leq k \leq N, |l - k| > 1 \] \[ \tilde{p}(s^k \mid s^k) := 1 - \tilde{p}(s^{k-1} \mid s^k) - \tilde{p}(s^{k+1} \mid s^k) \quad 1 \leq k \leq N - 1 \] \[ \tilde{p}(s^{N} \mid s^{N}) := 1 - \tilde{p}(s^{N-1} \mid s^{N}) , \]

where

\[ q^{UB,+}_k := \sum_{l=\bar{l}(k)+1}^{\mid C \mid} m_{c_\sigma(l)} \cdot \alpha_{c_\sigma(l)}^{\text{fail}} + \bar{m}(k) \cdot \alpha_{c_\sigma(\bar{l}(k))}^{\text{fail}} \quad 1 \leq k \leq N^{\text{max}} - 1 , \] \[ q^{UB,-}_k := \sum_{l=1}^{l(k)-1} m_{c_\sigma(l)} \cdot \alpha_{c_\sigma(l)}^{\text{repair}} + m(k) \cdot \alpha_{c_\sigma(l(k))}^{\text{repair}} \quad 1 \leq k \leq N^{\text{max}} . \] \[ 26 \text{ The idea of } \bar{l}(k) \text{ is that } \{ c_{\sigma(\bar{l}(k))}, \ldots, c_{\sigma(\mid C \mid)} \} \text{ is the smallest set of components with the largest failure rates that maximally require at least } N^{\text{max}} - k \text{ resources. The idea of } l(k) \text{ is that } \{ c_{\sigma(1)}, \ldots, c_{\sigma(l(k))} \} \text{ is the smallest set of components with the smallest repair rates that maximally require at least } \min\{ k, n_{\text{res}} \} \text{ resources. } \bar{l}(k), \bar{m}(k), l(k), \text{ and } m(k) \text{ are illustrated in Figure 5.9.} \]
Then, $\tilde{p}(\cdot \mid s^k)$, $1 \leq k \leq N$, are indeed probability distributions and the transition probabilities $\tilde{p}(s^l \mid s^k)$ satisfy the assumptions of Theorem 5.3.4.

In Theorem 5.3.6, $q_{UB,+}^k$ is the “maximum transition rate to a state with more failures” (from a state $x \in S^k$). $q_{UB,+}^k$ is given by the sum of the largest failure rates (where the failure rate of $c \in C$ is counted $m_c$ times). On the other hand, $q_{UB,-}^k$ is the “minimum transition rate to a state with less failures” (from a state $x \in S^k$). $q_{UB,-}^k$ is given by the sum of the smallest repair rates (where the repair rate of $c \in C$ is counted $m_c$ times). The idea of the summation of the failure/repair rates is illustrated in Figure 5.9. The proof of Theorem 5.3.6 is given in Appendix B.2 on page 228.

\[
\begin{array}{cccccc}
N_{\text{max}} - k & \overleftarrow{m(k)} & q_{UB,+}^k \\
\hline
\alpha_{f}^{\sigma(1)} & \alpha_{f}^{\sigma(2)} & \alpha_{f}^{l(k)} & \alpha_{f}^{(|C|)} \\
\hline
\alpha_{r}^{\sigma(1)} & \alpha_{r}^{\sigma(2)} & \alpha_{r}^{l(k)} & \alpha_{r}^{(|C|)} \\
\hline
\min\{k, n_{\text{res}}\} & \overrightarrow{m(k)} & q_{UB,-}^k
\end{array}
\]

**Figure 5.9** Illustration of $\tilde{l}(k)$, $\overline{m}(k)$, $\underline{l}(k)$, and $\underline{m}(k)$ as well as $q_{UB,+}^k$ and $q_{UB,-}^k$ of Theorem 5.3.6. In the figure, the notations $\alpha_i^f := \alpha_{c_{\sigma(i)}}^{\text{fail}}$ and $\alpha_i^r := \alpha_{c_{\sigma(i)}}^{\text{repair}}$ are used.
5.4 Case study

In this section we present results of a case study with data based on a real high-voltage power grid. The input data for this case study were kindly provided by RWE Rhein-Ruhr Netzserservice GmbH. In Section 5.4.1, the considered power grid and the problem instances are described. Results of the base model are shown in Section 5.4.2. Results of the aggregated and the base model are compared in Section 5.4.3. In Section 5.4.4, the aggregated model is applied to a larger problem instance, where the number of states prohibits the use of the base model. An approach to estimate the (neglected) effect of traveling is suggested in Section 5.4.5. The potential applicability of the model in a different context, where “calls for resources” occur much more often, is illustrated in Section 5.4.6. Finally, the case study is summarized in Section 5.4.7.

5.4.1 Power grid, input data, and problem instances

A schematic view of the considered high-voltage power grid is given in Figure 5.10. The power grid is operated at 110 kV and consists of 8 transformer substations (marked with \(V_1, V_2, \ldots, V_8\)) and approximately 253 km of overhead lines. The transformer substation \(V_1\) establishes the link to the extra-high-voltage power grid and the remaining 7 transformer substations \((V_2, \ldots, V_8)\) provide a connection to the medium-voltage power grid. Load points, which represent the demand from the medium-voltage power grid, are connected to the latter transformer substations. The demand of the 7 load points sums to 99 MW. The power grid covers an area of approximately 300 km\(^2\) and the load points aggregate the demand of approximately 31'000 consumers. The moderate size of the supply area justifies the neglect of travel times.

The power grid in Figure 5.10 consists of the following failing components \(c \in C\): 7 components corresponding to the 7 overhead lines (marked with \(l_1, l_2, \ldots, l_7\)), 6 common mode failures (affecting the lines \(l_1 + l_2, l_1 + l_5, l_3, l_6, l_6 + l_7, l_7\)), 8 components corresponding to the 8 busbars in HV (one in each of \(V_1, \ldots, V_8\)), 7 components corresponding to the 7 transformers HV/MV (one in each of \(V_2, \ldots, V_8\)) and 7 components corresponding to the 7 busbars in MV (one in each of \(V_2, \ldots, V_8\)). Throughout the case study, it is assumed that the transformer EHV/HV (in \(V_1\)) as well as the busbar in EHV (in \(V_1\)) do not fail.
Each failing component $c \in C$ is characterized by its failure rate $\alpha_{\text{fail}}^c$ and its repair rate $\alpha_{\text{repair}}^c$. Recall that $\alpha_{\text{fail}}^c$ and $\alpha_{\text{repair}}^c$ are the parameters of the exponential distributions of the failure and repair time, respectively, of (one identical parallel element of) $c$, cf. Section 5.1.2. Minimum and maximum values of the frequencies of failure (i.e., number of failures within a year) of the (identical parallel elements of the) failing components in this case study are provided in Table 5.1. The values differ for components of the same type due to their different characteristics, e.g., different number of associated switch bays. The frequencies of failure and the failure and repair rates are based on historical data.\footnote{cf. [Stö01, p. 149].} Details on the calculation of the failure and repair rates and the frequencies of failure are given in Appendix A.3 on page 186.

The mean time to failure (MTTF) and the mean time to repair (MTTR) of (one identical parallel element of) a component $c \in C$ are given by\footnote{cf. [BA92, p. 186, 282].}:

$$\text{MTTF}_c = \frac{1}{\alpha_{\text{fail}}^c},$$
$$\text{MTTR}_c = \frac{1}{\alpha_{\text{repair}}^c}.$$

For $c \in C^{\text{indep}}$, MTTF$_c$ is the expected time until the next independent
<table>
<thead>
<tr>
<th>Failure Type</th>
<th>Minimum (min)</th>
<th>Maximum (max)</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overhead lines (HV)</td>
<td>[0.0034, 0.0059]</td>
<td></td>
<td>km(^{-1}) a(^{-1})</td>
</tr>
<tr>
<td>Common mode failures</td>
<td>[0.0024, 0.0024]</td>
<td></td>
<td>km(^{-1}) a(^{-1})</td>
</tr>
<tr>
<td>Transformers HV/MV</td>
<td>[0.037, 0.037]</td>
<td></td>
<td>a(^{-1})</td>
</tr>
<tr>
<td>Busbars HV</td>
<td>[0.0010, 0.0084]</td>
<td></td>
<td>a(^{-1})</td>
</tr>
<tr>
<td>Busbars MV</td>
<td>[0.075, 0.11]</td>
<td></td>
<td>a(^{-1})</td>
</tr>
</tbody>
</table>

*Table 5.1* Minimum (min) and maximum (max) values of the frequencies of failure per km and year (km\(^{-1}\) a\(^{-1}\)) or per year (a\(^{-1}\)) of the failing components in this case study.

The total number of failing components is 35, 29 of which represent independent failures and thus belong to \(C^\text{indep}\) and 6 represent common mode failures and thus belong to \(C^\text{CM}\). 6 components in \(C^\text{indep}\) consist of only 1 element (namely, the power supply lines \(l_1, l_2, l_4, \text{ and } l_5\), and the transformers in \(V_3\) and \(V_4\)) and the remaining 23 components in \(C^\text{indep}\) consist of 2 identical parallel elements. Thus, the total number of states is

\[
|S| = 2^{12} \cdot 3^{23} = 3.86 \cdot 10^{14}.
\]

This huge number of states prohibits the calculation of the whole power grid model.
Table 5.2 Minimum (min) and maximum (max) values of the mean time to failure (MTTF) and the mean time to repair (MTTR) of the failing components in this case study.

<table>
<thead>
<tr>
<th>Component</th>
<th>MTTF [min, max]</th>
<th>MTTR [min, max]</th>
</tr>
</thead>
<tbody>
<tr>
<td>overhead lines (HV)</td>
<td>[3.5, 35]</td>
<td>[3.7, 5.2]</td>
</tr>
<tr>
<td>common mode failures</td>
<td>[11, 140]</td>
<td>[1.5, 1.5]</td>
</tr>
<tr>
<td>transformers HV/MV</td>
<td>[27, 27]</td>
<td>[140, 140]</td>
</tr>
<tr>
<td>busbars HV</td>
<td>[120, 960]</td>
<td>[5.5, 5.5]</td>
</tr>
<tr>
<td>busbars MV</td>
<td>[9.4, 13]</td>
<td>[9.6, 9.6]</td>
</tr>
</tbody>
</table>

Table 5.3 Summary of failing components corresponding to power supply lines (lines), common mode failures (CM failures), and busbars/transformers for problem instances $I_1$, $I_1^{\text{redund}(+)}$, $I_2$, and $I_2^{\text{redund}(+)}$. An entry $x+y$ for the number of components means $x$ components with 1 identical parallel element and $y$ components with 2 identical parallel elements.

<table>
<thead>
<tr>
<th>instance</th>
<th>number of failing components</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>lines</td>
</tr>
<tr>
<td>$I_1$</td>
<td>4+3</td>
</tr>
<tr>
<td>$I_1^{\text{redund}(+)}$</td>
<td>2+5</td>
</tr>
<tr>
<td>$I_2$</td>
<td>4+3</td>
</tr>
<tr>
<td>$I_2^{\text{redund}(+)}$</td>
<td>2+5</td>
</tr>
</tbody>
</table>

grid with the base model. As a consequence, we restrict ourselves to a subset of failing components for results of the base model and use the aggregated model to obtain results for the whole power grid.

We consider two problem instances $I_1$ and $I_2$ (and modifications thereof). In the first problem instance $I_1$, we consider failures of the power supply lines only and assume that neither the transformers nor the busbars fail. This leads to a total number of 13 failing components and the number of states is $|S_{I_1}| = 2^{10} \cdot 3^3 = 27'648$. In the second problem instance $I_2$, we consider all 35 failing components. An overview of the number of failing components for problem instances $I_1$ and $I_2$ (and of two modifications that are explained below) is given in Table 5.3.

We filter out all transient states and work with $S^{\text{recurrent}}$ (cf. Theorem 5.2.6 and Remark 5.2.8) and the discrete-time MDPs $\mathcal{M}^{\text{rec}}$ (cf. (5.15);
base model) and \( \tilde{\mathcal{M}}_K \) (cf. (5.25); aggregated model).\(^{30}\) An overview of the total number of states and the number of recurrent and transient states for problem instances \( I_1 \) and \( I_2 \) (and again of two modifications) is given in Table 5.4. Filtering out the transient states is particularly useful in the base model of instance \( I_1 \) (and \( I_1^{\text{redund}(+)} \)) as it allows to reduce the number of states by 64\% (and 77\%).

<table>
<thead>
<tr>
<th>instance ( I_1 )</th>
<th>model</th>
<th>recurrent</th>
<th>transient</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>base</td>
<td>9840</td>
<td>17'808</td>
<td>27'648</td>
<td></td>
</tr>
<tr>
<td>aggregated, ( K = 3 )</td>
<td>421</td>
<td>6</td>
<td>427</td>
<td></td>
</tr>
<tr>
<td>( I_1^{\text{redund}(+)} )</td>
<td>base</td>
<td>58'425</td>
<td>190'407</td>
<td>248'832</td>
</tr>
<tr>
<td>( I_2 )</td>
<td>base</td>
<td>73'345</td>
<td>204</td>
<td>73'549</td>
</tr>
<tr>
<td>aggregated, ( K = 4 )</td>
<td>8027</td>
<td>6</td>
<td>8033</td>
<td></td>
</tr>
<tr>
<td>aggregated, ( K = 3 )</td>
<td>92'645</td>
<td>290</td>
<td>92'935</td>
<td></td>
</tr>
</tbody>
</table>

**Table 5.4** Number of states for problem instances \( I_1 \), \( I_1^{\text{redund}(+)} \), \( I_2 \), and \( I_2^{\text{redund}(+)} \).

In Figure 5.11, the total number of controls (i.e., \( \sum_{x \in S_{\text{recurrent}}} |U(x)| \)) and the total number of possible transitions (i.e., \( \sum_{x \in S_{\text{recurrent}}} \sum_{u \in U(x)} \sum_{\bar{x} \in S_{\text{recurrent}}} \mathbf{1}_{\{p(\bar{x}|x,u) \neq 0\}} \)) for instance \( I_1 \) in the base model are depicted against the number of available resources \( n_{\text{res}} \). The maximum total number of controls is 271'541 and the maximum total number of possible transitions is 1'952'112 and both maxima are attained for \( n_{\text{res}} = 4 \).

Lower and upper bounds on the optimal average expected power not supplied \( \lambda_{\text{M}_{\text{rec}}}^* \) (or \( \lambda_{\text{M}_{\text{K}}}^* \)) are calculated with the relative value iteration algorithm as described in Section 4.2.3. The parameters that were used in the algorithm are summarized in Table 5.5.

As in Chapter 3, the continuity of supply is indirectly assessed by the non-availability of supply \( Q_{HUV} \) in minutes per year (measured by the ASIDI, cf. Section 3.5). \( Q_{HUV} \) is obtained from the optimal average expected power not supplied \( \lambda^* \) (where \( \lambda^* = \lambda_{\text{M}_{\text{rec}}}^* \) or \( \lambda^* = \lambda_{\text{M}_{\text{K}}}^* \)), which is

---

\(^{30}\) The uniformization parameter \( \eta_M \) is calculated with \( f_\eta := 1/0.9999 \), cf. (5.10).
5.4 Case study

Figure 5.11 Total number of controls and total number of possible transitions for instance $I_1$ in the base model.

<table>
<thead>
<tr>
<th>instance</th>
<th>$I_1$</th>
<th>$I_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>stopping accuracy [MW]</td>
<td>$\epsilon$</td>
<td>$10^{-10}$</td>
</tr>
<tr>
<td>reference state</td>
<td>$x_r$</td>
<td>$(0, \ldots, 0)$</td>
</tr>
</tbody>
</table>

Table 5.5 Parameters for the relative value iteration algorithm.

measured in MW, by

$$Q_U^{HV} = \frac{\lambda^* \cdot 365 \cdot 24 \cdot 60 \text{ min/year}}{L_{total}^{HV}},$$  \hspace{1cm} (5.39)

where $L_{total}^{HV}$, the total connected load, is taken to be the total demand in the system, i.e., $L_{total}^{HV} = 99$ MW for the considered power grid.

5.4.2 Results of base model for instance $I_1$

Figure 5.12(a) shows upper and lower bounds on the optimal non-availability of supply $Q_U^{HV}$ of problem instance $I_1$ against the number of available resources $n_{res}$, calculated with the base model of Section 5.2. Figure 5.12(b) shows the results starting from $n_{res} = 2$ to make the gap between upper and lower bounds visible. The non-availability of supply is approximately 8.15 minutes, independently of the number of resources. The improvement in the non-availability of supply from 1 to 2 resources amounts to less than 0.1 seconds per year. The results suggest that 1 resource is sufficient to operate instance $I_1$ in the given setting. Additional resources only lead to a marginal improvement in the non-availability of supply and thus can hardly be legitimated. In fact, with the chosen accu-
The insensitivity of the results with respect to the number of available resources is primarily explained by the very rare occurrence of failures and the redundancies in the power grid. Table 5.6 gives values of the stationary distribution corresponding to the \( \epsilon \)-optimal policy for the set of states \( S^0 \), which consists of only 1 state (the one where no failure has occurred), and the set of states \( S^1 \), which contains all states that require 1 resource. The values are valid for all numbers of available resources \( n_{\text{res}} \in \{1, 2, \ldots, 11\} \). The stationary distribution yields that the system is fully operating (i.e., has 0 failures) in more than 99.9% of the time. In less than 0.06% of the time, (at least) 1 resource is required. Due to the redundancies in the power grid, in only 7 states in \( S^1 \) (of totally 13), some load has to be shed. The histogram of the power not supplied for the states in \( S^1 \) is given in Figure 5.13. Finally, in less than 0.007% of the time, 2 or more resources are required. This result also reflects the \((n - 1)\) criterion (cf. Section 2.1.3), which requires that the grid can cope with a single failure.

<table>
<thead>
<tr>
<th>states</th>
<th>stationary distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S^0 )</td>
<td>0.99943 \ldots</td>
</tr>
<tr>
<td>( S^1 )</td>
<td>5.63 \ldots \cdot 10^{-4}</td>
</tr>
</tbody>
</table>

**Table 5.6** Values (non-rounded) of the stationary distribution corresponding to the \( \epsilon \)-optimal policy for being in a state of \( S^0 \) or \( S^1 \) for instance \( I_1 \) (valid for all \( n_{\text{res}} \in \{1, 2, \ldots, 11\} \)).
5.4 Case study

Modified variants of instance $I_1$

Next we present results of some modified variants of instance $I_1$. In the first set of modifications $I_1^f$, we multiply all failure rates by a factor $f \in \{0.5, 1.5, 2\}$. In Figure 5.14(a), the upper bounds on the optimal non-availability of supply of problem instances $I_1, I_1^{0.5}, I_1^{1.5},$ and $I_1^2$ are displayed against the number of available resources $n_{\text{res}}$. Figure 5.14(b) shows the results of variant $I_1^2$ starting from $n_{\text{res}} = 2$ including the lower bounds.

The non-availability of supply of the modified instances is approximately\textsuperscript{31} $f$ times the one of the “normal” variant $I_1$. In contrast to the normal variant (cf. Figure 5.12(b)), the effect on the non-availability of\textsuperscript{31} The non-availability of supply is not exactly scaled by $f$, because the repair rates were left unchanged.

\textsuperscript{31}
supply of a third resource can be quantified for $I_1 \times 2$. It seems that the additional resources become more valuable when the failure rates are increased. However, the improvement from 2 to 3 resources only amounts to approximately 0.06 milliseconds per year and is thus still negligible. For increased failure rates (with the considered factors), the effect of more than 1 resource is marginal and the failure rate has a much higher impact on the non-availability of supply. This hints at the importance of maintenance to keep the failure rates at the present level.

In the second set of modifications, we change the available redundancies in the power grid. In modification $I_1^{\text{redund}(-)}$ we remove the redundant (identical parallel) power supply lines in $l_3$, $l_6$, and $l_7$ and in modification $I_1^{\text{redund}(+)}$ we add a parallel power supply line in $l_1$ and $l_4$ and include two additional common mode failures, one that affects the (now two) lines of $l_1$ and one that affects the (now two) lines of $l_4$.\footnote{The failure rate of the additional common mode failure in $l_1$ is based on the part of the power supply line $l_1$ that is not affected by the common mode failures $l_1 + l_2$ and $l_1 + l_5$.} The idea of $I_1^{\text{redund}(+)}$ is to assure that no load is shed in case of any single independent failure, i.e., $I_1^{\text{redund}(+)}$ is in compliance with the $(n - 1)$ criterion. Of course, most of the common mode failures still lead to some load being shed. The number of components of $I_1^{\text{redund}(+)}$ is summarized in Table 5.3 and the number of (recurrent and transient) states in Table 5.4. Figure 5.15 depicts the upper bounds on the optimal non-availability of supply of problem instances $I_1$, $I_1^{\text{redund}(-)}$, and $I_1^{\text{redund}(+)}$ against the number of available resources $n_{\text{res}}$.

![Figure 5.15](image_url)

**Figure 5.15** Non-availability of supply for instances $I_1$, $I_1^{\text{redund}(-)}$, and $I_1^{\text{redund}(+)}$. The results show a large impact of the redundancies in the power grid. By
removing three parallel lines (in $I_1^{\text{redund}(-)}$) the non-availability of supply is almost doubled. On the other hand, by inserting two additional parallel lines and including two additional common mode failures (in $I_1^{\text{redund}(+)}$), the non-availability of supply is reduced by more than 65%.

**Influence of the optimal policy**

In the following, we analyze the influence of the choice of the policy on the non-availability of supply. To this end, we compare the optimal non-availability of supply with the worst possible non-availability of supply. The latter can be calculated with the relative value iteration algorithm\textsuperscript{33} by replacing “min” with “max” in step 2 of the algorithm (cf. Section 4.2.3).

Figure 5.16(a) shows the upper bounds on the optimal and the worst-case non-availability of supply of problem instance $I_1$ against the number of available resources $n_{\text{res}}$. Figure 5.16(b) shows the results starting from $n_{\text{res}} = 2$ and additionally shows the corresponding lower bounds.

Figure 5.16  Optimal and worst-case non-availability of supply for instance $I_1$ (UB: upper bound, LB: lower bound).

For the chosen accuracy, a difference in the optimal and the worst-case non-availability of supply is only observed for $n_{\text{res}} = 1$ (in (a)) and $n_{\text{res}} = 2$ (in (b)). The effect of an optimal policy on the non-availability of supply is approximately 0.5 seconds per year for $n_{\text{res}} = 1$ and only around 0.09 milliseconds per year for $n_{\text{res}} = 2$. Choosing a non-optimal policy thus only has a marginal effect on the resulting non-availability of supply. For states that require at most 1 resource (in which the system is during most

\textsuperscript{33} with the same parameters as before, cf. Table 5.5.
of the time), there is no choice of control (for $n_{\text{res}} \geq 1$). Thus the choice of the policy only has an effect for states that require at least 2 resources. As these states are hardly ever attained, the influence of the (optimal) policy on the non-availability of supply is very low.

5.4.3 Results of aggregated model for instance $I_1$

Next we compare the base model to the aggregated model of Section 5.3. Figure 5.17 shows the scatter plot of the power not supplied of each state against the corresponding number of required resources. The values on the vertical line corresponding to $k$ required resources thus represent the power not supplied of states in $S^k$. The plot shows that already in $S^2$, there is at least one state that attains the maximum power not supplied of 99 MW (and thus to obtain an upper bound on the non-availability of supply with the aggregated model, we have to set $\tilde{g}(s^k) = 99$ for $k \geq 2$, cf. (5.31), (5.32), (and (5.38)).

![Figure 5.17 Scatter plot of the power not supplied of all states of instance $I_1$ against the corresponding number of required resources. The red line shows the average power not supplied for each set of states $S^k$.](image)

In Figure 5.18, the (upper bounds on the) optimal non-availability of supply of problem instance $I_1$ for the aggregated and the base model are shown against the number of available resources $n_{\text{res}}$. For increasing $K$, the values of the aggregated model approach the values of the base model. Already for $K \in \{2, 3\}$, the non-availability of supply of the aggregated model is close to the one of the base model. Again, this can be explained by the fact that the system hardly ever visits states that require “many” resources and thus the states in $S^k$, $k > K$, for “large enough” $K$ can be aggregated. In the instance at hand, $K = 3$ seems to be sufficient.
On the other hand, for $K \in \{0, 1\}$, the aggregated model yields (very) pessimistic bounds on the non-availability of supply.

### 5.4.4 Results of aggregated model for instance $I_2$

We now show results of the second problem instance $I_2$ where all 35 failing components are considered. As stated above, the total number of states in $I_2$ prohibits calculations with the base model and the aggregated model has to be used. In the following we discuss effects that are observed in the aggregated model. It is important to remark that it cannot be concluded that the effects would be of comparable magnitude in the base model. However, by Theorem 5.3.4, the upper bound on the optimal non-availability of supply of the aggregated model always provides an upper bound on the optimal non-availability of supply of the base model.
Figure 5.19 shows upper and lower bounds on the optimal non-availability of supply of the aggregated model with $K \in \{3, 4\}$ for problem instance $I_2$ against the number of available resources $n_{\text{res}}$.

The non-availability of supply is approximately 47 minutes per year for all numbers of available resources. Again, as for the smaller instance $I_1$, the impact of a second (or third, etc.) resource is rather negligible (the second resource decreases the non-availability of supply only by approximately 13 seconds per year for $K = 4$). Compared to the real non-availability of supply observed in reality, the 47 minutes per year are far too high. The reason for this is that instance $I_2$ does not satisfy the $(n - 1)$ criterion. Namely, the two HV/MV transformers in $V_3$ and $V_4$ and – as already discussed for instance $I_1$ – the power supply lines $l_1$ and $l_4$ have no redundant element. In practice, should one of these components fail, there is the possibility to resupply the corresponding consumers via the medium-voltage power grid. However, this effect is not captured in our model. Therefore, we consider the modified problem instance $I_2^{\text{redund}(+)}$, where two identical parallel transformers are installed in $V_3$ as well as in $V_4$ and the power supply lines and common mode failures are modified as for $I_1^{\text{redund}(+)}$. The number of components of $I_2^{\text{redund}(+)}$ is summarized in Table 5.3 and the number of (recurrent and transient) states in Table 5.4. Upper and lower bounds on the optimal non-availability of supply for $I_2^{\text{redund}(+)}$ against the number of available resources $n_{\text{res}}$ are depicted in Figure 5.20 for the aggregated model with $K = 4$. For $I_2^{\text{redund}(+)}$ a more realistic non-availability of supply of approximately 3 minutes per year is obtained. We can thus assign an approximate value of 44 minutes of non-availability of supply per year to the additionally installed elements. The additional redundant elements allow to model approximately the
possibility of resupplying consumers via the medium-voltage power grid. For $I_2^{\text{redund}(+)}$, the second resource again decreases the non-availability of supply by approximately 13 seconds per year which, in this case, is an improvement of around 7%.

**Influence of the optimal policy**

Figure 5.21 shows the upper and lower bounds on the optimal and the worst-case non-availability of supply of the aggregated model with $K = 4$ for problem instance $I_2^{\text{redund}(+)}$ against the number of available resources $n_{res}$. For $n_{res} = 1$, a non-optimal policy leads to an increase of up to 2.55 minutes per year (+80%) in the non-availability of supply, i.e., the policy has a significant influence on the non-availability of supply for $n_{res} = 1$. For $n_{res} = 2$, the difference in the non-availability of supply between an
optimal and a worst-case policy is less than 0.5 seconds per year. As there is a significant difference in the non-availability of supply for \( n_{\text{res}} = 1 \), the system has to visit states that require two or more resources "sufficiently often", because an optimal policy can only make a difference in these states. This is confirmed by the stationary distribution corresponding to the \( \epsilon \)-optimal policy, cf. Table 5.7. For \( n_{\text{res}} = 1 \), states in \( S^2 \) are visited significantly more often than for \( n_{\text{res}} = 2 \).

<table>
<thead>
<tr>
<th>states</th>
<th>stationary distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S^0 )</td>
<td>0.989212... 0.989267...</td>
</tr>
<tr>
<td>( S^1 )</td>
<td>0.010642... 0.010677...</td>
</tr>
<tr>
<td>( S^2 )</td>
<td>0.000143... 0.000054...</td>
</tr>
</tbody>
</table>

**Table 5.7** Values (non-rounded) of the stationary distribution corresponding to the \( \epsilon \)-optimal policy for being in a state of \( S^0 \), \( S^1 \), or \( S^2 \) for instance \( I_2^{\text{redund}(+)} \).

The discussed effects are observed in the aggregated model and it cannot be concluded that the same effects would be observed in the base model. However, by optimizing the policy in the aggregated model for \( n_{\text{res}} = 1 \), a significantly stronger upper bound on the optimal non-availability of supply of the base model is obtained.

**Increased demand for instance \( I_2^{\text{redund}(+)} \)**

In the following we consider a different demand scenario. Namely, the demand of the load point in \( V_i \), \( i \in \{2, \ldots, 8\} \), is set to 50% of the total capacity of all transformers in \( V_i \). In this case, the total demand is 165 MW and thus \( L_{\text{total}}^{HV} = 165 \) MW for the calculation of \( Q_{hv}^{U} \) (cf. (5.39)). Figure 5.22 shows upper and lower bounds on the optimal non-availability of supply of the aggregated model with \( K = 4 \) for the normal and the increased demand scenario for problem instance \( I_2^{\text{redund}(+)} \) against the number of available resources \( n_{\text{res}} \). Compared to the normal demand scenario (with 99 MW), the non-availability of supply for the increased demand scenario is only approximately 0.3 minutes per year worse (for all considered numbers of resources). The effect of a second resource in the increased demand scenario is approximately equal to the improvement in the normal scenario. It is to remark that the optimal average expected power not supplied \( (\lambda_{\text{opt}}^{\star M_K}) \) is approximately 1.8 times higher
5.4 Case study

2.8
2.9
3.0
3.1
3.2
3.3
3.4
3.5
3.6
1 2 3 4
non-availability of supply [min/year]
number of resources
normal demand, UB
normal demand, LB
increased demand, UB
increased demand, LB

Figure 5.22  Non-availability of supply of the aggregated model with $K = 4$ for instance $I^\text{redund(+)}_2$ for the normal and the increased demand scenario (UB: upper bound, LB: lower bound).

for the increased demand (for all considered numbers of resources). The small difference in the non-availability of supply is thus also an effect of the calculation with a different $L^\text{HV}_{\text{total}}$.

5.4.5 Effect of traveling

In the present model, travel times of the resources are neglected. The rather small size of the power grid in this case study implies that the resources that are responsible for this power grid are in fact responsible for a larger power grid. Thus, in case of a failure, the resources might be outside the considered power grid and the travel time to reach the failed component might not be negligible. To simulate this effect, we modify the mean time to repair of all failing components by adding a fixed repair offset $\tau$, i.e., the modified repair rates $\alpha^\text{repair,}\tau_c$ satisfy $\frac{1}{\alpha^\text{repair,}\tau_c} = \frac{1}{\alpha^\text{repair}_c} + \tau$ (for all $c \in C$). The repair offset $\tau$ is intended to model the (maximum) average time a resource needs to travel to the site of a failure. In Figure 5.23, upper and lower bounds on the optimal non-availability of supply of the aggregated model with $K = 4$ are depicted for problem instance $I^\text{redund(+)}_2$ and for repair offsets $\tau = 0$ hours (normal case) and $\tau \in \{1, 2, 10\}$ hours. The repair offset has a significant influence on the non-availability of supply. For instance, a repair offset of 2 hours more than doubles the non-availability of supply. On the other hand, the influence of a second resource remains small, even for a repair offset of 10 hours. Again, this is mainly explained by the high reliability of the power grid. Even when the repair takes significantly longer, the
probability of two or more simultaneous failures is very small. However, if the resources are responsible for a larger power grid, there is a second effect in addition to the increased travel times. Namely, some resources might already be working outside the considered power grid and thus are temporarily unavailable. To obtain more realistic results, this additional effect would have to be incorporated.

5.4.6 More sensitive results for an (unrealistic) instance $I_1 \times 10^6$

For the sake of illustration, we consider the modification $I_1 \times 10^6$ of the first instance $I_1$ where the failure rates are multiplied with a factor of $10^6$. Upper and lower bounds of the resulting optimal non-availability of supply against the number of available resources $n_{\text{res}}$ for the base model and the aggregated model with $K \in \{6, 7, 8\}$ are shown in Figure 5.24. In this case, the non-availability of supply reacts very sensitively to the number of available resources. As failures occur much more often, the system requires considerably more resources and thus each additional resource leads to a significant decrease in the non-availability of supply. In contrast to the previous (realistic) cases, the aggregation threshold $K$ has to be chosen larger to obtain “close” bounds. Even though the considered instance is not at all realistic in the context of power grids, the results show the potential applicability of the described model in a different context, where “calls for resources” (failures) occur more often.
5.4 Case study

![Figure 5.24](image)

**Figure 5.24** Non-availability of supply for instance $I_1^{1,000}$ for the base model and the aggregated model with $K \in \{6, 7, 8\}$ (UB: upper bound, LB: lower bound).

### 5.4.7 Summary

In the case study we investigated the effects of a limited number of resources on the non-availability of supply in the base model and the aggregated model for different problem instances. Apparently, the non-availability of supply is determined much more by the redundancies in the power grid and the failure/repair rates than the number of available resources. The results show that the aggregated model yields close upper bounds already for small aggregation thresholds $K$. The main reason for these findings is that failures are very rare and thus (almost) always at most one resource is required. The repair duration of a failure is mainly driven by the (given) repair rate of the corresponding component and there is hardly any influence of the availability of resources because the resources are almost always “idle” and thus available. Finally, even for increased repair times, which are used to model the traveling of the resources, the effect of additional resources on the non-availability of supply is small.

It is also shown that the possibility of resupplying consumers via the underlying MV power grid (which is not included in the model) can be approximated by considering additional redundant elements.

In the largest considered instance ($I_2^{\text{redud}(+)}$), there is a significant influence of the assignment decision for $n_{\text{res}} = 1$ (in the aggregated model). Namely, a non-optimal policy can increase the non-availability of supply by up to 80%. For $n_{\text{res}} = 1$, the system thus requires two or more resources during a “significant” amount of time because an optimal policy can only make a difference in these situations.
Comparing the results of the first and the second model, the influence of the availability of resources on the continuity of supply is much stronger in the first model, in particular in MV/LV. In the considered MV/LV grids in the first model, each failure results in an interruption of supply and requires resources on site for resupplying the affected consumers. On the other hand, due to the redundancies in the HV grid of the second model, consumers are hardly ever affected by failures. The more redundancies there are and the higher the degree of automation and remote control in a power grid, the less crucial is the immediate availability of resources for the repair work. This also hints at the necessity of a more detailed and sophisticated model for the power grid to obtain more realistic results.
Chapter 6

Conclusions
6.1 Conclusions

The restoration time of an incident in the power grid, i.e., the duration from the incident’s occurrence until its full repair, strongly depends on the availability of the human resources performing the restoration work. To take strategic decisions concerning the organization of these resources, power grid operators need to quantify the effects of a limited number of resources and their organization on the restoration times and finally on the continuity of supply, which is one main aspect of the quality of supply. In this dissertation, two models are presented that allow the analysis and quantification of these effects. In both models, the restoration times endogenously depend on the current availability of a limited number of resources.

The first model simulates the activities of the resources for the restoration after incidents in all voltage levels. By incorporating many organizational details, such as the time-dependent availability of resources, and by mimicking the real processes of the restoration after incidents, the model allows a detailed analysis of the restoration times and in particular of delays due to the traveling or the limited availability of the resources. The continuity of supply can be assessed by a variety of key performance indicators that are evaluated by the model. In a case study based on a real supply area, the effects of different organizations of resources on the continuity of supply are analyzed and compared. The results suggest, for instance, that a reduction in the number of resources during periods of standby service can be partly compensated by collaborating resources, either between different areas of responsibility or different voltage levels. Further, with the new key performance indicators that are based on incidents without interruption of supply, the performance of an organization of resources in power grids with redundancies, where most incidents are without interruption of supply, can be evaluated.

In the first model, the effects of incidents without interruption of supply are assessed by their so-called power-at-risk. Being a static measure, this value only approximates the effects of simultaneous incidents in redundant power grids. This simplification in the first model led to the idea for the second model, whose focus lies on the “vulnerability” of redundant power grids over time. The power grid is modeled by independent components with exponential failure and repair rates, where the repair can only be performed if a resource is available. Thus, in contrast to traditional reliability evaluations, the restoration times depend on the availability
of a limited number of resources. Compared to state-of-the-art reliability evaluations, however, the power grid model in the second model is very simplified in many aspects. The second model is formulated as a continuous-time Markov decision process that minimizes the long-term average expected power not supplied and is solved via its discrete-time counterpart. To reduce the large number of states, an approximate model with aggregated states is formulated that provides an upper bound on the base model. In a case study for a real high-voltage power grid, the effects of a limited number of resources on the continuity of supply are quantified and analyzed. The results do not show any unexpected effects but much more confirm the high reliability of redundant power grids. Due to the redundancies and the rare occurrence of incidents, the continuity of supply reacts very insensitively to the number of resources. Additional or fewer redundancies as well as changes in the failure rates have a much higher impact on the continuity of supply than the number of available resources. Nevertheless, the effect of additional resources can be explicitly measured with the model. Due to the high reliability of the power grid, the aggregated model yields close upper bounds even for a high level of aggregation. Hence with the aggregated model, the size of the considered problem instances can be increased.

In principle, the type of effects of a limited number of resources on the continuity of supply is predictable. The more resources there are and the less restricted they operate, the better is the continuity of supply. These intuitive dependencies are verified by both models. The strength of the models, however, lies in their capability of actually quantifying the extent of these effects. Consequently, different organizations of resources can objectively be evaluated and compared, thus supporting the process of decision making.

The effect of the availability of resources on the continuity of supply is significantly larger in the first model than in the second model. This is related to the available redundancies and the degree of automation and remote control in the considered power grids. In MV/LV, incidents usually infer an interruption of supply and resources are required on site to re-establish the supply (first model). In redundantly operated grids with automation and remote control, incidents generally do not lead to an interruption of supply and the repair work does not directly affect the continuity of supply (second model). It can be concluded that the fewer redundancies and the lower the degree of automation and remote control in a power grid, the more important is the precise modeling of
the availability of the resources to quantify the effects of an organization of resources on the continuity of supply. In contrast, in power grids with many redundancies and a high degree of automation and remote control, the immediate availability of resources for repair is less crucial and the detailed modeling of the power grid becomes more significant.

Both models also reveal some of the strengths and limitations of mathematical modeling. Due to the simulative nature of the first model, real-world problem instances can be computed. The many operational details, however, prohibit the direct optimization of the assignment or the organization of resources. In the second model, the simplified operational processes allow the optimization of the assignment of resources to incidents. On the other hand, the exponentially growing number of states limits the size of the power grids that can be considered and the approximate model has to be used for larger problem instances.

The development of the first model, which was done in close collaboration with RWE Rhein-Ruhr Netzservice GmbH, is an exemplar of the field of Operations Research. First, the real problem of the grid service company of finding an (optimal) balance between cost and quality of supply, was analyzed and abstracted into a conceptual model. The conceptual relations were then transformed into a mathematical model. This model was applied to real data and the results were finally interpreted in the context of the real problem. The good quality of the results confirm that the model captures the relevant aspects of the real situation. The successful use of the model and its results by RWE Rhein-Ruhr Netzservice GmbH for supporting the solution of the real problem further adds to the quality and practical relevance of the model.

6.2 Outlook

Finally, we provide some ideas for further investigations. In both models, the assignment of the resources to incidents is instantaneous and the coordination processes in the control center are neglected. Operating with a limited number of resources, the control center might also face shortages of resources that affect the coordination and thus the duration of the restoration processes. The first model could thus be extended by incorporating the processes in the control center.

The coordination of the resources becomes particularly critical during an extreme situation with many simultaneous incidents. The assignment
model of the first model, adapted to an online setting, might provide decision support for the control center in such cases.

One main drawback of the second model is the exponentially growing state space. By redefining the notion of a “failing component”, e.g., to represent a whole part of the power grid, larger grids could be considered. At some point, however, travel times of the resources would have to be incorporated.

The power grid models in state-of-the-art reliability evaluation models are far more detailed and sophisticated than the one used in the second model. It would be desirable to extend these models with an endogenous dependency of the restoration times on a limited availability of resources, as proposed in the second model.

The applicability of the models in a different context could also be investigated. The setting of a limited number of “resources” that has to accomplish randomly arriving “tasks” applies to many other applications not related to power grids. In the first model, the assignment of resources to tasks is based on the characterization of each single task by its “impact” and the (estimated) “processing time”. The assignment in the second model is based on the integrated impact of all currently present tasks on the system. By appropriately interpreting these characteristic values, the models of this thesis might be used in different fields.

Finally, the first model could be extended by a power grid model that, as in the second model, dynamically determines the integral effect of simultaneous incidents. This information could then be used in the assignment of resources to incidents.
Appendix A

Additional material
In this first part of the appendix, some additional material is presented. Some additional properties of discrete-time homogeneous Markov chains are given in Section A.1. Section A.2 contains further remarks on the DC load shedding model of Section 5.2.5. Finally, Section A.3 describes the estimation of the failure and repair rates for the case study in Section 5.4.

### A.1 More on discrete-time homogeneous Markov chains

In this section, we provide some additional properties of discrete-time homogeneous Markov chains (DTHMCs). We use the same notations as in Section 4.1.1, but – whenever possible – prove the statements for a countable state space. We begin with elementary properties of accessible states. Using (4.2), we conclude for distinct states \( i, j \in S, i \neq j \), that

\[
i \rightarrow j \iff p(i_1 \mid i_0)p(i_2 \mid i_1) \cdots p(i_n \mid i_{n-1}) > 0 \tag{A.1}
\]

for states \( i_0, i_1, \ldots, i_n \in S \) with \( i_0 := i \) and \( i_n := j \).

Hence we have

\[
(i \to j \text{ and } j \to k) \Rightarrow i \to k, \tag{A.2}
\]

for all \( i, j, k \in S \).

The following theorem provides a characterization of closed sets.

**Theorem A.1.1** A set \( C \subseteq S \) is closed if and only if

\[
(i \in C \text{ and } i \to j) \Rightarrow j \in C. \tag{A.3}
\]

**Proof of Theorem A.1.1:**

“\( \Rightarrow \)” Let \( i \in C \) and \( j \in S \) with \( i \to j \). We have to show that \( j \in C \).

If \( j = i \), this is trivially true. We thus assume \( j \neq i \) and suppose that \( j \notin C \). By definition of \( i \to j \), there is some \( n \in \mathbb{N}_0 \) such that \( p^{(n)}(j \mid i) > 0 \). As \( j \neq i \), \( n \neq 0 \). If \( n = 1 \), \( p^{(1)}(j \mid i) = p(j \mid i) > 0 \)

\[\text{1 cf. [Nor97, Theorem 1.2.1].}\]
\[\text{2 cf. [Nor97, p. 11].}\]
\[\text{3 [Nor97, p. 11] defines closed communicating classes via (A.3).}\]
which, as \( j \not\in C \), is a contradiction to \( C \) being closed. Hence, \( n \geq 2 \), and by the Chapman-Kolmogorov equations (4.1), we have

\[
0 < p^{(n)}(j \mid i) = \sum_{k \in S} p^{(1)}(k \mid i) \cdot p^{(n-1)}(j \mid k)
\]

\[(*) \sum_{k \in C} p^{(1)}(k \mid i) \cdot p^{(n-1)}(j \mid k)\]

where we used in \((*)\) that \( p(k \mid i) = 0 \) for all \( k \not\in C \). Thus, there is a \( k \in C \) such that \( p^{(n-1)}(j \mid k) > 0 \). Repeating this argument, we finally obtain some \( l \in C \) with \( p^{(1)}(j \mid l) = p(j \mid l) > 0 \), a contradiction to \( C \) being closed.

“⇐” Let \( i \in C \), \( j \not\in C \). We have to show that \( p(j \mid i) = 0 \). Due to (A.3), \( i \not\rightarrow j \), i.e., \( p^{(n)}(j \mid i) = 0 \) for all \( n \in \mathbb{N}_0 \), and thus \( p^{(1)}(j \mid i) = p(j \mid i) = 0 \).

\[\square\]

The next theorem relates the structural properties of DTHMCs that allow the same transitions.\(^4\)

**Theorem A.1.2** Let \( \mathcal{X} \) and \( \hat{\mathcal{X}} \) denote two DTHMCs on the (countable) set of states \( S \) with transition probabilities \( p(j \mid i) \) and \( \hat{p}(j \mid i) \), \( i, j \in S \), respectively. We assume that for distinct states \( i, j \in S \), \( i \neq j \),

\[
p(j \mid i) = 0 \iff \hat{p}(j \mid i) = 0.
\]

Then for \( i, j \in S \) we have

\[
i \rightarrow j \text{ for } \mathcal{X} \iff i \rightarrow j \text{ for } \hat{\mathcal{X}}. \tag{A.4}
\]

Furthermore, if \( |S| < \infty \), we have for \( i \in S \)

(i) \( i \) is recurrent for \( \mathcal{X} \iff i \) is recurrent for \( \hat{\mathcal{X}} \),

(ii) \( i \) is transient for \( \mathcal{X} \iff i \) is transient for \( \hat{\mathcal{X}} \).

**Remark A.1.3** Without the assumption \( |S| < \infty \), (i) and (ii) of the above theorem do not hold in general. Consider for example the simple random walk\(^5\) \( \mathcal{X} = \{X_n \mid n \in \mathbb{N}_0\} \) on \( \mathbb{Z} \). \( \mathcal{X} \) is defined by \( X_0 := 0 \)

\[^{4}\text{cf. also [Ber07, p. 315, 316] in the context of semi-Markov decision processes.}\]

\[^{5}\text{cf. [Res92, p. 33], [Nor97, Example 1.6.1].}\]
and \( X_n := \sum_{i=1}^{n} Y_i, \ n \geq 1, \) where the \( Y_i \) are independent and identically distributed with \( \mathbb{P}[Y_i = 1] = p \) and \( \mathbb{P}[Y_i = -1] = 1 - p \) for some \( 0 < p < 1 \). Two simple random walks with different parameters \( p \) satisfy the assumption of Theorem A.1.2, but the simple random walk is recurrent if \( p = \frac{1}{2} \) and transient otherwise.\(^6\)

**Proof of Theorem A.1.2:** As, by definition, \( p^{(0)}(i \mid i) = \hat{p}^{(0)}(i \mid i) = 1 \) for all \( i \in S \), we have \( i \rightarrow i \) for all \( i \in S \) and (A.4) holds for \( j = i \). Let \( i, j \in S \) with \( j \neq i \). (A.1) together with the assumption on the transition probabilities yields (A.4) for \( j \neq i \).

Proof of (ii), “\( \Leftarrow \)”. Let \( S = T \cup C_1 \cup C_2 \cup \cdots \cup C_K \) be the unique partition of the state space into transient states \( T \) and closed recurrent communicating classes \( C_k, 1 \leq k \leq K \), corresponding to the chain \( \mathcal{X} \) (cf. (4.5)). Let \( i \in S \) be a transient state of \( \hat{\mathcal{X}} \) and assume that \( i \) is recurrent for \( \mathcal{X} \). Then \( i \in C_k \) for some \( 1 \leq k \leq K \). Due to (A.4) and (A.3), \( C_k \) is a closed communicating class of \( \hat{\mathcal{X}} \). By [Nor97, Theorem 1.5.6], every finite closed communicating class is recurrent. As \( S \) is finite, \( C_k \) is finite and thus recurrent for \( \hat{\mathcal{X}} \), which is a contradiction. Replacing \( \mathcal{X} \) and \( \hat{\mathcal{X}} \) in the above line of argument proves “\( \Rightarrow \)” of (ii). Finally, (i) follows from (ii) by contraposition. \( \Box \)

The following theorem provides a useful property of the stationary distribution of a DTHMC. Thinking of probability as flow,\(^7\) the theorem states that the total flow out of a set of states equals the total flow into this set of states (cf. Figure A.1).\(^8\)

![Figure A.1](image-url)  
**Figure A.1** Total “flow” out of \( S' \) equals total “flow” into \( S' \).

---

\(^6\) cf. [Nor97, Example 1.6.1].  
\(^7\) cf. [KS95].  
\(^8\) The theorem was inspired by the “flow equations” in [MdSeSG89, Proof of Lemma 2].
Theorem A.1.4 Let $\mathcal{X}$ denote a DTHMC and assume that $\mathcal{X}$ has a stationary distribution $\pi$. Let $S' \subseteq S$. Then we have (cf. Figure A.1)

$$\sum_{i \in S'} \pi_i \sum_{j \in S \setminus S'} p(j \mid i) = \sum_{i \in S \setminus S'} \pi_i \sum_{j \in S'} p(j \mid i).$$

Proof of Theorem A.1.4: For each $i \in S$, we have $1 = \sum_{j \in S} p(j \mid i) = \sum_{j \in S'} p(j \mid i) + \sum_{j \in S \setminus S'} p(j \mid i)$. Define $\pi(S') := \sum_{i \in S'} \pi_i$. Then

$$\pi(S') = \sum_{i \in S'} \pi_i \cdot 1$$

$$= \sum_{i \in S'} \pi_i \cdot \left( \sum_{j \in S'} p(j \mid i) + \sum_{j \in S \setminus S'} p(j \mid i) \right)$$

$$= \sum_{i \in S'} \pi_i \sum_{j \in S'} p(j \mid i) + \sum_{i \in S'} \pi_i \sum_{j \in S \setminus S'} p(j \mid i). \quad (A.5)$$

Using the condition $\pi_j = \sum_{i \in S} \pi_i p(j \mid i)$ on the stationary distribution (for all $j \in S$), we obtain

$$\pi(S') = \sum_{j \in S'} \pi_j$$

$$= \sum_{j \in S'} \sum_{i \in S} \pi_i p(j \mid i)$$

$$= \sum_{j \in S'} \left( \sum_{i \in S'} \pi_i p(j \mid i) + \sum_{i \in S \setminus S'} \pi_i p(j \mid i) \right)$$

$$= \sum_{i \in S'} \pi_i \sum_{j \in S'} p(j \mid i) + \sum_{i \in S \setminus S'} \pi_i \sum_{j \in S'} p(j \mid i). \quad (A.6)$$

Combining (A.5) and (A.6) yields the result. □
The next theorem shows how to aggregate the states (called microstates) of a DTHMC into so-called macrostates such that the stationary probability of a macrostate equals the sum of the (original) stationary probabilities of the corresponding microstates (cf. Figure A.2).\(^9\)

\[ Q_{kl} := \frac{\sum_{i \in A_k} \pi_i \sum_{j \in A_l} p(j | i)}{\sum_{i \in A_k} \pi_i} \quad (A.7) \]

Then, Q is a stochastic matrix and the aggregated DTHMC \(X^{agg}\) on the state space \(\{1, \ldots, N\}\) with transition matrix Q is irreducible and the stationary distribution \(\pi_Q\) of \(X^{agg}\) satisfies

\[ (\pi_Q)_k = \sum_{i \in A_k} \pi_i \quad 1 \leq k \leq N. \]

\(^9\) See [Mey89, Theorem 2.2, Theorem 2.3, Theorem 4.1 (and its proof)], using that the transition matrix of a DTHMC is irreducible (cf. [BP79, Definition 2.1.2]) if and only if the DTHMC is irreducible [BP79, Theorem 8.3.9]. [KS95, Proposition 1] reformulate the result as we state it.

\(^{10}\) cf. Theorem 4.1.4.
A.2 Further remarks on the DC load shedding model in Section 5.2.5

In this section we make some further remarks on the DC load shedding model of Section 5.2.5. The following theorem characterizes the solutions of

\[ p = B(x)\theta \]  

(cf. (5.20)). For the notations, we refer to Section 5.2.5. For simplicity we assume that the graph \((V,E^{\text{on}}(x))\) is connected. However, the results of Theorem A.2.1 (appropriately modified) hold for any connected subgraph of \((V,E^{\text{on}}(x))\) (cf. Remark 5.2.9(ii)).

**Theorem A.2.1** Let \(x \in S\) and assume that the graph \((V,E^{\text{on}}(x))\) is connected. Let \(p \in \mathbb{R}^{|V|}\) be given.

(i) If (A.8) has a solution \(\theta\), then this solution is unique up to an additive constant.

(ii) (A.8) has a solution \(\theta\) if and only if \(\sum_{v=1}^{|V|} p_v = 0\).

(iii) \(\text{rank}(B(x)) = |V| - 1\).

(iv) Assume \(\sum_{v=1}^{|V|} p_v = 0\). Let \(v, w \in \{1, \ldots, |V|\}\) and \(\alpha \in \mathbb{R}\). Let \(e_v \in \mathbb{R}^{|V|}\) denote the \(v\)-th unit vector, let \(B_{\backslash w}(x)\) denote the submatrix of \(B(x)\) after removing the \(w\)-th row, and let \(p_{\backslash w}\) denote the subvector of \(p\) after removing the \(w\)-th element. Then

\[ \theta := \left( B_{\backslash w}(x) e_v^T \right)^{-1} \cdot \begin{pmatrix} p_{\backslash w} \\ \alpha \end{pmatrix} \]  

(A.9)

is the unique solution of (A.8) with \(\theta_v = \alpha\) (and the inverse exists).

---

**Proof of Theorem A.2.1:** For notational convenience, we omit the dependence on \(x\) and we write \(E^{\text{on}}, r_{\{v,w\}}, B,\) and \(B_{\backslash w}\) instead of \(E^{\text{on}}(x), r_{\{v,w\}}(x), B(x),\) and \(B_{\backslash w}(x),\) respectively.

(i) Choose a direction on all line segments (edges) \(\{v, w\} \in E^{\text{on}}\) and let \(\overrightarrow{E}\) denote the set of directed edges. For a solution \(\theta\) of (A.8) and a directed edge \(\overrightarrow{e} = (v, w) \in \overrightarrow{E}\) (directed from \(v\) to \(w\)), define

\[ f_{\overrightarrow{e}} := \frac{\theta_v - \theta_w}{r_{\{v,w\}}} \]  

(A.10)
i.e., $f_{\overrightarrow{v}}$ is the power flow from $v$ to $w$, cf. (5.21). Let $A \in \{-1, 0, 1\}^{V | E}$ with columns $A_{\overrightarrow{e}}$, $\overrightarrow{e} = (v, w) \in \overrightarrow{E}$, defined by

$$
(A_{\overrightarrow{e}})_k := \begin{cases} 
+1 & \text{if } k = v \\
-1 & \text{if } k = w \\
0 & \text{otherwise},
\end{cases}
$$

i.e., $-A$ is the incidence matrix\(^{11}\) of the graph $(V, \overrightarrow{E})$. Thus, (A.10) can be written as

$$
\theta^T A_{\overrightarrow{e}} = f_{\overrightarrow{e}} \cdot r_{\{v, w\}}.
$$

(A.11)

First we show that

$$
\sum_{\overrightarrow{e} \in \overrightarrow{E}} A_{\overrightarrow{e}} f_{\overrightarrow{e}} = B \theta.
$$

(A.12)

The $\bar{v}$-th component of $\sum_{\overrightarrow{e} \in \overrightarrow{E}} A_{\overrightarrow{e}} f_{\overrightarrow{e}}$ is given by

$$
\sum_{\overrightarrow{e} \in \overrightarrow{E}} (A_{\overrightarrow{e}})_{\bar{e}} f_{\overrightarrow{e}} = \sum_{\overrightarrow{e} = (v, w) \in \overrightarrow{E}} (A_{\overrightarrow{e}})_{\bar{e}} \cdot \frac{\theta_v - \theta_w}{r_{\{v, w\}}} \\
= \sum_{\overrightarrow{e} : \overrightarrow{e} = (\bar{v}, w) \in \overrightarrow{E}} (+1) \cdot \frac{\theta_\bar{v} - \theta_w}{r_{\{\bar{v}, w\}}} + \sum_{\overrightarrow{e} : \overrightarrow{e} = (w, \bar{v}) \in \overrightarrow{E}} (-1) \cdot \frac{\theta_w - \theta_\bar{v}}{r_{\{w, \bar{v}\}}} \\
= \sum_{w : \{\bar{v}, w\} \in E_{on}} \frac{\theta_\bar{v} - \theta_w}{r_{\{\bar{v}, w\}}} \\
= \theta_\bar{v} \sum_{w : \{\bar{v}, w\} \in E_{on}} \frac{1}{r_{\{\bar{v}, w\}}} - \sum_{w : \{\bar{v}, w\} \in E_{on}} \theta_w \frac{1}{r_{\{\bar{v}, w\}}} \\
= \theta_\bar{v} B_{\bar{v}\bar{v}} + \sum_{w : \{\bar{v}, w\} \in E_{on}} \theta_w B_{\bar{v}w} \\
= \sum_{w \in V} B_{\bar{v}w} \theta_w = (B \theta)_\bar{v}.
$$

Now let $\theta, \hat{\theta}$ be two solutions of (A.8) and let $f_{\overrightarrow{e}}, \hat{f}_{\overrightarrow{e}}, \overrightarrow{e} \in \overrightarrow{E}$, be

\(^{11}\) cf. [Sch03, p. 35].
defined by (A.10). As $B\theta = B\hat{\theta} = p$, we have

$$
\sum_{\vec{e} \in \bar{E}} A_{\vec{e}}(f_{\vec{e}} - \hat{f}_{\vec{e}}) = \sum_{\vec{e} \in \bar{E}} A_{\vec{e}} f_{\vec{e}} - \sum_{\vec{e} \in \bar{E}} A_{\vec{e}} \hat{f}_{\vec{e}} \tag{A.12}
$$

By multiplying the last sequence of equations with $(\theta - \hat{\theta})^T$ from the left, we obtain

$$
0 = \sum_{\vec{e} \in \bar{E}} (\theta - \hat{\theta})^T A_{\vec{e}}(f_{\vec{e}} - \hat{f}_{\vec{e}})
= \sum_{\vec{e} \in \bar{E}} \left( \theta^T A_{\vec{e}} r_{\{v,w\}} - \hat{\theta}^T A_{\vec{e}} r_{\{v,w\}} \right)(f_{\vec{e}} - \hat{f}_{\vec{e}})
= \sum_{\vec{e} \in \bar{E}} (f_{\vec{e}} - \hat{f}_{\vec{e}})^2 \cdot r_{\{v,w\}}
$$

from which we conclude (as $r_{\{v,w\}} > 0$) $f_{\vec{e}} - \hat{f}_{\vec{e}} = 0$ for all $\vec{e} \in \bar{E}$.

Hence $\theta v - \theta w = \hat{\theta} v - \hat{\theta} w$ for all $\vec{e} = (v, w) \in \bar{E}$ or, equivalently, $\theta v - \hat{\theta} v = \theta w - \hat{\theta} w$ for all $\vec{e} = (v, w) \in \bar{E}$. As the graph is connected, we thus have $\theta v - \hat{\theta} v = \alpha \forall v$ (for some $\alpha \in \mathbb{R}$) and we have shown that if $\hat{\theta}$ is a solution of (A.8), we have

$$
\{ \theta | B\theta = p \} = \{ \hat{\theta} + \alpha e | \alpha \in \mathbb{R} \} \tag{A.13}
$$

with $e := (1, \ldots, 1)^T \in \mathbb{R}^{|V|}$.

(ii) By the orthogonal decomposition theorem [Mey00, p. 405], we have $\text{image}(B) = (\text{kernel}(B^T))^\perp$, i.e., $p \in \text{image}(B) \iff p^T \lambda = 0$ for all $\lambda \in \text{kernel}(B^T)$. Hence

$$
\exists \theta : B\theta = p \iff \exists \lambda : \lambda^T B = 0, \lambda^T p \neq 0. \tag{A.14}
$$

As $B = B^T$, we have $\{ \lambda | \lambda^T B = 0 \} = \{ \lambda | B\lambda = 0 \}$ and as $0$ is a solution of (A.8) for $p = 0$, we obtain from (A.13)

$$
\{ \lambda | \lambda^T B = 0 \} = \{ \alpha e | \alpha \in \mathbb{R} \}.
$$

Thus, (A.14) can be written as

$$
\exists \theta : B\theta = p \iff \exists \alpha : \alpha e^T p \neq 0
\iff \alpha e^T p = 0 \forall \alpha.
$$

But the last statement is true if and only if $e^T p = 0$. 
(iii) For $p = 0$, $\theta = 0$ is a solution of (A.8). From (A.13) we conclude that the kernel of $B$ has dimension 1. Thus, $\text{rank}(B) = |V| - 1$.\(^{12}\)

(iv) As $e^T B = 0$ and as we assume $e^T p = 0$, we can remove an arbitrary equation of $B\theta = p$ without changing the set of solutions. By setting $\theta_v := \alpha$, we conclude from (i) and (ii) that the system of equations

$$
\begin{pmatrix}
B \setminus w \\
e_v^T
\end{pmatrix} \theta = \begin{pmatrix} p \setminus w \\ \alpha \end{pmatrix}
$$

has a unique solution (which is a solution of (A.8)). Thus, the (quadratic) matrix in the above equation has rank $n$ and is invertible.

□

In the sequel we discuss a (somewhat artificial) condition that assures that $\kappa(\cdot)$ is monotone (cf. Remark 5.2.12). Let $E_{\text{aux,trafo}} \subseteq E$ denote the auxiliary line segments that connect transformers to busbars. Recall that these line segments are used to incorporate the capacity constraints of the transformers. Further let $V_{\text{trafo}}^{\text{EHV}} \subseteq V_{\text{trafo}}$ denote the subset of transformers that transform power from EHV to HV, i.e., these transformers are connected to feed-in points and provide power to the high-voltage power grid. The remaining transformers $V_{\text{trafo}} \setminus V_{\text{trafo}}^{\text{EHV}}$ thus transform power from HV to MV. Finally, let $V_{\text{load}}^{\text{MV}} \subseteq V_{\text{load}}$ denote the set of load points in MV. The remaining load points $V_{\text{load}} \setminus V_{\text{load}}^{\text{MV}}$ are directly connected to busbars in HV. We make the following assumptions:

Each transformer $v \in V_{\text{trafo}}^{\text{EHV}}$ is connected to its own feed-in point (via its own busbar in EHV).  \hspace{1cm} (A.15)

Each transformer $v \in V_{\text{trafo}} \setminus V_{\text{trafo}}^{\text{EHV}}$ is connected to its own load point (via its own busbar in MV).  \hspace{1cm} (A.16)

Assumptions (A.15) and (A.16) avoid that transformers with different capacities that are connected to the same feed-in point or load point limit each other in the DC load shedding model. Due to assumption (A.15), each feed-in point $v \in V_{\text{feed}}$ feeds power to a unique transformer $v'(v) \in V_{\text{trafo}}^{\text{EHV}}$. Thus the capacity limit $p_{v'(v)}^\text{max}(x)$ of $v'(v)$ (in state $x \in S$) leads to the constraint

$$
p_v \leq p_{v'(v)}^\text{max}(x) \hspace{1cm} (A.17)
$$

\(^{12}\) By the rank plus nullity theorem [Mey00, p.199]: $\dim(\ker(B)) + \rank(B) = |V|$. 
on the net real power of \( v \). Due to assumption (A.16), each load point \( v \in V_{\text{load}}^{\text{MV}} \) is supplied by a unique transformer \( v'(v) \in V_{\text{trafo}} \setminus V_{\text{EHV}}^{\text{trafo}} \). For \( v \in V_{\text{load}} \) and each state \( x \in S \), define

\[
p_v^{\text{possible}}(x) := \begin{cases} 
\min \{ p_{v'}^{\text{max}}(x), p_v^{\text{load}} \} & v \in V_{\text{load}}^{\text{MV}} \\
p_v^{\text{load}} & v \in V_{\text{load}} \setminus V_{\text{load}}^{\text{MV}}.
\end{cases}
\]

For \( v \in V_{\text{load}}^{\text{MV}} \), \( p_v^{\text{possible}}(x) \) gives the maximum amount of power that can be supplied to \( v \) by the corresponding transformer. Hence the net real power of \( v \in V_{\text{load}} \) has to satisfy (for all \( x \in S \))

\[
-p_v \leq p_v^{\text{possible}}(x).
\]

(A.18)

**Theorem A.2.2** Assume (A.15) and (A.16) hold. If

\[
c_{\{v,w\}}^{\text{single}} \geq \sum_{v' \in V_{\text{load}}} p_{v'}^{\text{load}} \quad \text{for all } \{v, w\} \in E \setminus E_{\text{aux,trafo}}
\]

(A.19)

i.e., if each single identical line segment in \( E \setminus E_{\text{aux,trafo}} \) can carry the total demand in the system, then \( \kappa(\cdot) \) is monotone.

If the capacities of the line segments in \( E \setminus E_{\text{aux,trafo}} \) are more restrictive, \( \kappa(\cdot) \) is not necessarily monotone, as shown in the example in Remark 5.2.12. Assumption (A.19) of the previous theorem implies that the capacity constraints for lines in \( E \setminus E_{\text{aux,trafo}} \) in the LP (5.22) are always satisfied because no (feasible) power flow on any line segment can be larger than the total demand in the system. Under assumptions (A.15), (A.16), and (A.19), the LP (5.22) reduces to (using (A.17) and (A.18))

\[
\kappa(x) = \min \sum_{v \in V_{\text{load}}} (p_v^{\text{load}} + p_v)
\]

(A.20)

s.t. \( p = B(x)\theta \)

\[
0 \leq p_v \leq p_{v'(v)}^{\text{max}}(x) \quad v \in V_{\text{feed}}
\]

\[
-p_v^{\text{possible}}(x) \leq p_v \leq 0 \quad v \in V_{\text{load}}
\]

\[
p_v = 0 \quad v \in V_{\text{busbar}} \cup V_{\text{trafo}} \cup V_{\text{aux}}
\]

\[
\theta_v \in \mathbb{R} \quad v \in V.
\]
Proof of Theorem A.2.2: Let \( x, x' \in S \) with \( x \leq x' \), i.e., \( x_c \leq x'_c \) for all \( c \in C \). Let \((\tilde{V}_1(x), \tilde{E}_1(x)), \ldots, (\tilde{V}_L(x), \tilde{E}_L(x))\) denote the connected subgraphs of \((V,E_{\text{on}}(x))\) where \( 1 \leq L \leq |V| \). Either \( x' \) induces the same subgraphs as \( x \) or \( x' \) induces at least one additional subgraph. However, each additional subgraph induced by \( x' \) is a subgraph of some \((\tilde{V}_l(x), \tilde{E}_l(x))\), \( 1 \leq l \leq L \). For \( 1 \leq l \leq L \), let \((\tilde{V}_{l,i}(x'), \tilde{E}_{l,i}(x'))\), \( 1 \leq i \leq n_l \in \mathbb{N} \), denote the subgraphs of \((\tilde{V}_l(x), \tilde{E}_l(x))\) induced by \( x' \), i.e., the set of subgraphs induced by \( x' \) is \( \{(\tilde{V}_{l,i}(x'), \tilde{E}_{l,i}(x')) \mid 1 \leq l \leq L, 1 \leq i \leq n_l \}\).

Let \((p^*, \theta^*)\) denote an optimal solution of LP (A.20) corresponding to state \( x' \). Recall from Remark 5.2.9(ii) that \( B(x') \) has a block structure with an independent block for each connected subgraph induced by \( x' \). Let \( B_{l,i}(x') \) denote the submatrix (block) of \( B(x') \) corresponding to the node set \( \tilde{V}_{l,i}(x') \). Due to the definition of \( B(x') \) we have 
\[
(1, \ldots, 1) B_{l,i}(x') = (0, \ldots, 0)
\]
and thus \( p^* = B(x') \theta^* \)
\[
\sum_{v \in \tilde{V}_{l,i}(x')} p^*_v = 0 \quad 1 \leq l \leq L, 1 \leq i \leq n_l.
\]

Hence for each connected subgraph \((\tilde{V}_l(x), \tilde{E}_l(x))\), \( 1 \leq l \leq L \), induced by \( x \), we have
\[
\sum_{v \in \tilde{V}_l(x)} p^*_v = \sum_{i=1}^{n_l} \sum_{v \in \tilde{V}_{l,i}(x')} p^*_v = 0.
\]

For each connected subgraph \((\tilde{V}_l(x), \tilde{E}_l(x))\), \( 1 \leq l \leq L \), let \( \hat{\theta}^l = (\hat{\theta}^l_v)_{v \in \tilde{V}_l(x)} \) be defined as in (A.9) (for some \( v, w \in \tilde{V}_l(x) \) and some \( \alpha \in \mathbb{R} \)). Let \( \hat{\theta} = (\hat{\theta}_v)_{v \in V} \) such that \( \hat{\theta}_v = \hat{\theta}^l_v \) for all \( v \in \tilde{V}_l(x) \), \( 1 \leq l \leq L \). As a consequence, we have \( p^* = B(x) \hat{\theta} \).

Let \( v' \in V_{\text{trafo}} \) be a transformer. As \( x' \geq x \), we have \( n_{v'}(x') \leq n_{v'}(x) \) and thus
\[
p^\max_{v'}(x') \leq p^\max_{v'}(x) \quad v' \in V_{\text{trafo}}.
\] (A.21)

Recall that for \( v \in V_{\text{feed}} \), \( v'(v) \in V_{\text{trafo}}^{\text{EHV}} \) is the unique transformer that is fed by \( v \). From (A.21), we have
\[
p^\max_{v'(v)}(x') \leq p^\max_{v'(v)}(x) \quad v \in V_{\text{feed}}.
\] (A.22)
Recall that for \( v \in V_{\text{load}} \), \( v'(v) \in V_{\text{trafo}} \setminus V_{\text{trafo}}^{\text{EHV}} \) is the (unique) transformer which supplies \( v \). Due to (A.21), we have \( p_v^{\text{possible}}(x') \leq p_v^{\text{possible}}(x) \) for \( v \in V_{\text{load}}^{\text{MV}} \). As \( p_v^{\text{possible}}(x') = p_v^{\text{possible}}(x) \) for \( v \in V_{\text{load}} \setminus V_{\text{load}}^{\text{MV}} \), we have

\[
-p_v^{\text{possible}}(x) \leq -p_v^{\text{possible}}(x') \quad v \in V_{\text{load}}. \tag{A.23}
\]

Due to (A.22) and (A.23), \((p^*, \bar{\theta})\) is a feasible solution for LP (A.20) corresponding to state \( x \) and thus \( \kappa(x) \leq \sum_{v \in V_{\text{load}}} (p^\text{load}_v + p^*_v) = \kappa(x') \).

\[\square\]

Under assumptions (A.15), (A.16), and (A.19), the minimum power not supplied for a state \( x \in S \) is the sum of the lack of available capacity in each connected subgraph. This is formalized in the following theorem.

**Theorem A.2.3** Assume (A.15), (A.16), and (A.19) hold. Let \( x \in S \) be a state of the system and let \((\tilde{V}_1, \tilde{E}_1), \ldots, (\tilde{V}_L, \tilde{E}_L)\) denote the connected subgraphs of \((V, E^\text{on}(x))\) where \( 1 \leq L \leq |V| \).\(^{13}\) Then the minimum power not supplied is (where for \( a \in \mathbb{R} \), \( a^+ := \max\{a, 0\} \))

\[
\kappa(x) = \sum_{l=1}^{L} \left( \sum_{v \in \tilde{V}_l \cap V_{\text{load}}} (p^\text{load}_v - p_v^{\text{possible}}(x)) + \left( \sum_{v \in \tilde{V}_l \cap V_{\text{load}}} p_v^{\text{possible}}(x) - \sum_{v \in \tilde{V}_l \cap V_{\text{trafo}}^{\text{EHV}}} p_v^{\text{max}}(x) \right)^+ \right). \tag{A.24}
\]

**Proof of Theorem A.2.3:** The minimum power not supplied \( \kappa(x) \) in the system is the sum of the minimum power not supplied in each connected subgraph \( l \), \( 1 \leq l \leq L \), because each connected subgraph \( l \) can be considered independently (cf. Remark 5.2.9(ii)). For a connected subgraph \( l \), let \( \kappa_l(x) \) denote the minimum power not supplied in this subgraph. \( \kappa_l(x) \) is at least as large as the lack of available capacity in

\(^{13}\) For notational convenience, we omit the dependency of \((\tilde{V}_l, \tilde{E}_l), 1 \leq l \leq L, \) on \( x \).
this subgraph, i.e.,

$$
\kappa_l(x) \geq \sum_{v \in \tilde{V} \cap V_{\text{load}}} (p_{v}^{\text{load}} - p_{v}^{\text{possible}}(x)) + \left( \sum_{v \in \tilde{V} \cap V_{\text{load}}} p_{v}^{\text{possible}}(x) - \sum_{v \in \tilde{V} \cap V_{\text{EHV trafo}}} p_{v}^{\text{max}}(x) \right)^+ \quad (A.25)
$$

The first addend accounts for a potential lack of capacity in HV/MV transformers and the second addend accounts for a potential lack of capacity in EHV/HV transformers. If enough “feed-in” capacity is available in subgraph \( l \), i.e., if

$$
\sum_{v \in \tilde{V} \cap V_{\text{load}}} p_{v}^{\text{possible}}(x) \leq \sum_{v \in \tilde{V} \cap V_{\text{EHV trafo}}} p_{v}^{\text{max}}(x),
$$

let \( p_v := -p_v^{\text{possible}}(x) \) for \( v \in \tilde{V} \cap V_{\text{load}} \) (cf. (A.18)), and choose the \( p_v, v \in \tilde{V} \cap V_{\text{feed}}, \) such that \( p_v \leq p_v^{\text{max}}(x) \) (cf. (A.17)) and such that \( \sum_{v \in \tilde{V}} p_v = 0 \). Then, \( (p, \theta) \) with \( \theta \) as in (A.9) (for some \( v, w \in \tilde{V} \) and some \( \alpha \in \mathbb{R} \)) is a feasible solution of the LP (A.20) whose objective function value is

$$
\sum_{v \in \tilde{V} \cap V_{\text{load}}} (p_{v}^{\text{load}} + p_{v}) = \sum_{v \in \tilde{V} \cap V_{\text{load}}} (p_{v}^{\text{load}} - p_{v}^{\text{possible}}(x)),
$$

which equals the value on the right hand side of inequality (A.25) in this case. Thus, this solution is optimal.

On the other hand, if there is a lack of “feed-in” capacity in subgraph \( l \), i.e., if

$$
\sum_{v \in \tilde{V} \cap V_{\text{load}}} p_{v}^{\text{possible}}(x) > \sum_{v \in \tilde{V} \cap V_{\text{EHV trafo}}} p_{v}^{\text{max}}(x),
$$

let \( p_v := p_v^{\text{max}}(v)(x) \) for \( v \in \tilde{V} \cap V_{\text{feed}} \) (cf. (A.17)), and choose the \( p_v, v \in \tilde{V} \cap V_{\text{load}}, \) such that \( -p_v \leq p_v^{\text{possible}}(x) \) (cf. (A.18)) and such that \( \sum_{v \in \tilde{V}} p_v = 0 \). Then, \( (p, \theta) \) with \( \theta \) as in (A.9) (for some \( v, w \in \tilde{V} \) and some \( \alpha \in \mathbb{R} \)) is a feasible solution of the LP (A.20). Observe

$$
\sum_{v \in \tilde{V} \cap V_{\text{load}}} p_v = -\sum_{v \in \tilde{V} \cap V_{\text{feed}}} p_v = -\sum_{v \in \tilde{V} \cap V_{\text{feed}}} p_v^{\text{max}}(v)(x) = -\sum_{v \in \tilde{V} \cap V_{\text{EHV trafo}}} p_{v}^{\text{max}}(x),
$$
i.e., the objective function value of \((p, \theta)\) is
\[
\sum_{v \in \tilde{V} \cap V_{\text{load}}} (p_v^{\text{load}} + p_v) = \sum_{v \in \tilde{V} \cap V_{\text{load}}} p_v^{\text{load}} - \sum_{v \in \tilde{V} \cap V_{\text{trabo}}} p_v^{\max}(x),
\]
which equals the value on the right hand side of inequality (A.25) in this case. Thus, this solution is optimal. \qed
A.3 Calculation of failure and repair rates

In this section, we describe how we estimate the failure and repair rates $\alpha_{\text{fail}}^c$ and $\alpha_{\text{repair}}^c$ of the failing components $c \in C$ in the case study of Section 5.4. We first describe the case of $c \in C^{\text{indep}}$. Recall from Section 5.1.2 that for $c \in C^{\text{indep}}$, $\alpha_{\text{fail}}^c$ and $\alpha_{\text{repair}}^c$ are the parameters of the exponential distributions of the failure and repair time, respectively, of independent failures of one identical parallel element of $c$. For $c \in C^{\text{indep}}$, let $e_c$ denote such an identical parallel element of $c$. $\alpha_{\text{fail}}^c$ and $\alpha_{\text{repair}}^c$ are given by\(^{14}\)

\[
\alpha_{\text{fail}}^c = \frac{\text{number of failures of } e_c \text{ in a given time period}}{\text{time during which } e_c \text{ was operating in this period}},
\]

and

\[
\alpha_{\text{repair}}^c = \frac{\text{number of repairs of } e_c \text{ in a given time period}}{\text{total duration of these repairs}}.
\]

To each element $e_c$ we associate subelements that contribute to failures of $e_c$. For instance, to a transformer from HV to MV we associate one switch bay in HV and one switch bay in MV and failures of these switch bays also contribute to failures of the transformer. To overhead lines we associate one switch bay for each endpoint and the number of failures depend on the length of the line. If (a part of) an overhead line is not affected by common mode failures, we further associate an appropriate “amount” of common mode failures to the line. For each subelement $j$, $\hat{H}_j$, the estimated expected frequency of failure of $j$ within one year, and $\hat{T}_j$, the estimated mean time to repair of $j$ in hours are given from historical data.\(^{15}\) Denoting $J_{e_c}$ the set of subelements associated with $e_c$, we thus have

\[
\alpha_{\text{fail}}^c := \frac{\sum_{j \in J_{e_c}} \hat{H}_j}{8760 \text{h/a} - \sum_{j \in J_{e_c}} \hat{H}_j \cdot \hat{T}_j},
\]

\[
\alpha_{\text{repair}}^c := \frac{\sum_{j \in J_{e_c}} \hat{H}_j}{\sum_{j \in J_{e_c}} \hat{H}_j \cdot \hat{T}_j}.
\]

---

\(^{14}\) cf. [BA92, p. 282].

\(^{15}\) cf. [Stö01, p. 149].
A.3 Calculation of failure and repair rates

The estimated expected frequency of failure of $e_c$ within one year is $\hat{H}_{ec} := \sum_{j \in J_{ec}} \hat{H}_j$. As the estimated mean time to repair of $e_c$, $\hat{T}_{ec}$, is given by

$$\hat{T}_{ec} = \frac{1}{\alpha_{c}^{\text{repair}}} = \frac{\sum_{j \in J_{ec}} \hat{H}_j \cdot \hat{T}_j}{\hat{H}_{ec}},$$

we can write

$$\alpha_{c}^{\text{fail}} = \frac{\hat{H}_{ec}}{8760 \text{ h/a} - \hat{H}_{ec} \cdot \hat{T}_{ec}}.$$

For $c \in C^{\text{CM}}$, $\alpha_{c}^{\text{fail}}$ and $\alpha_{c}^{\text{repair}}$ are the parameters of the exponential distributions of the failure and repair time, respectively, of common mode failures of $c$ (cf. Section 5.1.2). From historical data, $\hat{H}_c$, the estimated expected frequency of common failures of $c$ within one year, and $\hat{T}^{\text{CM}}$, the estimated mean time to repair of a common mode failure, are given. Similarly to the case $c \in C^{\text{indep}}$, we have for $c \in C^{\text{CM}}$:

$$\alpha_{c}^{\text{fail}} := \frac{\hat{H}_c}{8760 \text{ h/a} - \hat{H}_c \cdot \hat{T}^{\text{CM}}},$$

$$\alpha_{c}^{\text{repair}} := \frac{1}{\hat{T}^{\text{CM}}}.$$
In this part of the appendix, the proofs that are mentioned in the main text are provided. For convenience of reading, the theorems are stated again.

B.1 Proofs of Chapter 3

Theorem 3.4.1 (on page 38) Let $\sigma$ be a permutation of $\{1,2,\ldots,|I(\omega)|\}$ that specifies the order in which the restoration for the incidents in $I(\omega)$ is performed. Then $\sigma$ is an optimal order (i.e., leads to the minimum in (3.4)) for the restricted special case if and only if $\sigma$ satisfies

\[
\frac{P_{1}^{\sigma(j)}}{\delta_{a}(j) + \delta_{s}(j)} \geq \frac{P_{1}^{\sigma(j+1)}}{\delta_{a}(j+1) + \delta_{s}(j+1)} \quad 1 \leq j \leq |I(\omega)| - 1 .
\]  

(B.1)

Proof of Theorem 3.4.1: The following proof is based on [Zen06]. Let $n := |I(\omega)|$ and let $\sigma$ be a permutation of $\{1,2,\ldots,n\}$ that specifies the order in which the restoration for the incidents is performed. The points in time $t_{\sigma(j)}$ when the resource starts to work on incident $\sigma(j)$ are given by

\[
t_{\sigma(1)} = 0 \\
t_{\sigma(j)} = t_{\sigma(j-1)} + \delta_{a}(j-1) + \delta_{s}(j-1) \quad 2 \leq j \leq n .
\]

Observe that minimizing (3.4) is equivalent to minimizing (over $\sigma$) the shaded area in Figure B.1. Minimizing the shaded area is again equivalent to minimizing (over $\sigma$) the area below the continuous, piece-wise linear function $f^{\sigma}(\cdot)$ that is indicated in Figure B.1. For $1 \leq j \leq n$ define

\[v_{j}^{\sigma} := -\frac{P_{1}^{\sigma(j)}}{\delta_{a}(j) + \delta_{s}(j)}, \text{ i.e., } v_{j}^{\sigma}, 1 \leq j \leq n - 1, \text{ equals the slope of } f^{\sigma}(\cdot) \text{ between } t^{\sigma(j)} \text{ and } t^{\sigma(j+1)} \text{ and } v_{n}^{\sigma} \text{ equals the slope of } f^{\sigma}(\cdot) \text{ between } t^{\sigma(n)} \text{ and } \bar{t} := t^{\sigma(n)} + \delta_{a}(n) + \delta_{s}(n).\]

Now suppose $\sigma$ does not satisfy (B.1), i.e., there exists a $k$, $1 \leq k \leq n-1$, such that $v_{k}^{\sigma} > v_{k+1}^{\sigma}$. In Figure B.1, this is illustrated for $k = 2$. Define $\sigma'$

---

\[\text{We identify each incident } j \in I(\omega) \text{ with an integer in } \{1,2,\ldots,|I(\omega)|\}.\]

\[\text{For the ease of notation, we write } t_{\sigma(j)} \text{ instead of } t_{1}^{\sigma(j)} \text{ (cf. Figure 3.1).}\]
as the permutation which is obtained from $\sigma$ by exchanging $k$ and $k + 1$. Then, the area below $f^{\sigma'}(\cdot)$ is smaller than the area below $f^{\sigma}(\cdot)$, i.e., $\sigma$ cannot be optimal. Hence, any optimal permutation satisfies (B.1).

Let $\sigma^*$ be an optimal permutation with area $g^*$ below $f^{\sigma^*}(\cdot)$. As was just proven, $\sigma^*$ satisfies (B.1). Now, if the permutation $\sigma$ also satisfies (B.1), the area below $f^{\sigma}(\cdot)$ equals $g^*$, i.e., $\sigma$ is also optimal.  \[ \square \]
B.2 Proofs of Chapter 5

Theorem 5.2.6 (on page 124) Let

\[ S_{\text{recurrent}} := \{ x \in S \mid C^{\text{CM}} \cap C^{\text{failed}}(x) = \emptyset \} \]

\[ \cup \{ x \in S \mid C^{\text{CM}} \cap C^{\text{failed}}(x) \neq \emptyset, x \text{ satisfies (A)} \} \]

where:

(A) There exists a permutation \( \sigma \) of the set \( \{1, \ldots, |C'|\} \) with

\[ \{ C_{\sigma(1)}, \ldots, C_{\sigma(|C^{\text{CM}} \cap C^{\text{failed}}(x)|)} \} = C^{\text{CM}} \cap C^{\text{failed}}(x) \]

such that (we use the convention that \( \bigcup_{i=a}^{b} \cdots = \emptyset \) whenever \( b < a \))

\[ C_{c_{\sigma(j)}}(x) \setminus \bigcup_{k=1}^{j-1} C_{c_{\sigma(k)}}(x) \neq \emptyset \quad 1 \leq j \leq |C^{\text{CM}} \cap C^{\text{failed}}(x)|. \]

Let \( n_{\text{res}} \geq 1 \) and let \( \mu \in \Pi^{\text{MD}} \) be a deterministic stationary policy. Then, for the embedded Markov chain \( \tilde{X}_\mu \) (of \( \tilde{\mathcal{M}} \)) corresponding to \( \mu \):

(i) \( S_{\text{recurrent}} \) is a closed recurrent communicating class.

(ii) All states in \( S_{\text{transient}} := S \setminus S_{\text{recurrent}} \) are transient.

Proof of Theorem 5.2.6 (details): The overview of the proof is given on page 126. We prove (1) and (2) in the following Lemma B.2.1 and Lemma B.2.2, respectively. Recall that for a deterministic stationary policy \( \mu \in \Pi^{\text{MD}}, \tilde{X}_\mu := \{ \tilde{X}_{\mu,n} \mid n \in \mathbb{N}_0 \} \) denotes the embedded Markov chain (of \( \tilde{\mathcal{M}} \)) corresponding to \( \mu \).

Lemma B.2.1 Let \( n_{\text{res}} \geq 1, x_0 \in S, \bar{x} \in S_{\text{recurrent}}, \) and \( \mu \in \Pi^{\text{MD}} \) a deterministic stationary policy. Then there exists \( N \in \mathbb{N}_0 \) such that

\[ \mathbb{P}[\tilde{X}_{\mu,N} = \bar{x} \mid \tilde{X}_{\mu,0} = x_0, \mu] > 0, \]

i.e., \( \bar{x} \) is accessible from \( x_0 \) (for \( \tilde{X}_\mu \)).

---

3 In the characterization of the states with (A) we explicitly use Assumption (5.1) which guarantees that \( C_c((0, \ldots, 0)) = C_c \neq \emptyset \) for all \( c \in C^{\text{CM}} \).

4 The idea of (A) is that the common mode failures that have occurred in \( x \) (corresponding to the components \( C^{\text{CM}} \cap C^{\text{failed}}(x) \)) can be ordered such that this order gives a sequence of failure events that can lead to state \( x \) with positive probability.
Proof of Lemma B.2.1:

(a) We first show that the state $x^0 = (0, \ldots, 0)$ is accessible from $x_0$, i.e., there exists $N' \in \mathbb{N}_0$ such that

$$\mathbb{P}[\tilde{X}_{\mu,N'} = x^0 \mid \tilde{X}_{\mu,0} = x_0, \mu] > 0. \quad (B.2)$$

If $x_0 = x^0$, (B.2) is true for $N' = 0$. We thus assume in the following that $x_0 \neq x^0$. As $x_0 \neq x^0$, some (independent or common mode) failures have occurred in $x_0$ and $C^\text{failed}(x_0) \neq \emptyset$. The control $\mu(x_0)$ thus assigns (at least) one resource to a component in $C^\text{failed}(x_0)$, cf. (5.5) on page 113. Let $x_1 \in S$ denote the state that is obtained from $x_0$ when this repair is finished. Then

$$\mathbb{P}[\tilde{X}_{\mu,1} = x_1 \mid \tilde{X}_{\mu,0} = x_0, \mu] = \tilde{p}(x_1 \mid x_0, \mu(x_0)) > 0. \quad (B.3)$$

If $x_1 = x^0$, (B.2) is true for $N' = 1$. We thus assume in the sequel that $x_1 \neq x^0$. As before, $\mu(x_1)$ assigns (at least) one resource to a component in $C^\text{failed}(x_1)$. Let $x_2 \in S$ denote the state that is obtained from $x_1$ when this repair is finished. Then, by the Chapman-Kolmogorov equations (4.1) we have

$$\mathbb{P}[\tilde{X}_{\mu,2} = x_2 \mid \tilde{X}_{\mu,0} = x_0, \mu] = \sum_{x \in S} \tilde{p}(x_2 \mid x, \mu(x))\tilde{p}(x \mid x_0, \mu(x_0))$$

$$\geq \tilde{p}(x_2 \mid x_1, \mu(x_1))\tilde{p}(x_1 \mid x_0, \mu(x_0)) \quad (4.19),(5.9)$$

$$> 0.$$

By repeating this argument, all components $c \in C^\text{CM} \cap C^\text{failed}(x_0)$ can be repaired (each having required one resource), and all components $c \in C^\text{indep} \cap C^\text{failed}(x_0)$ can be (fully) repaired (each having required at most $m_c$ resources) after a finite number of state transitions with positive probability, i.e., (B.2) holds for some $N' \in \mathbb{N}_0$.

(b) Next we show that $\bar{x}$ is accessible from $x^0$, i.e., $x^0 \rightarrow \bar{x}$. In other words, starting from $x^0$, all failures that are present in $\bar{x}$ (and only those) can occur in a finite number of state transitions with positive probability.

If $\bar{x} = x^0$, we clearly have $x^0 \rightarrow \bar{x}$. We thus assume in the sequel that $\bar{x} \neq x^0$ and thus $C^\text{failed}(\bar{x}) \neq \emptyset$. 


• Let $\bar{x}' \in S$ be component-wise defined as
  
  $$\bar{x}'_c := \begin{cases} 
  \bar{x}_c & \text{if } c \in C_{\text{failed}}(\bar{x}) \cap C_{\text{indep}} \\
  0 & \text{otherwise,}
  \end{cases}$$

  i.e., in $\bar{x}'$, the same independent failures have occurred as in $\bar{x}$ (and only these). We now argue that

  $$x^0 \rightarrow \bar{x}' . \quad (B.4)$$

  If $C_{\text{failed}}(\bar{x}) \cap C_{\text{indep}} = \emptyset$, $\bar{x}' = x^0$ and $x^0 \rightarrow \bar{x}'$ in this case. If $C_{\text{failed}}(\bar{x}) \cap C_{\text{indep}} \neq \emptyset$, choose an (arbitrary) order of the corresponding independent failures and let $x'_0, x'_1, \ldots, x'_n$ denote a sequence of states such that $x'_0 := x^0$, $x'_n := \bar{x}'$, and $x'_k$, $1 \leq k \leq n$, is the state which is obtained from $x'_{k-1}$ when the $k$-th independent failure (according to the chosen order) occurs.\(^5\) We have

  $$\tilde{p}(x'_k \mid x'_{k-1}, \mu(x'_{k-1})) \quad (4.19),(5.9) > 0 \quad 1 \leq k \leq n . \quad (B.5)$$

  Using (B.5), we conclude from (A.1) that $x^0 \rightarrow \bar{x}'$.

• Next we show that

  $$\bar{x}' \rightarrow \bar{x} . \quad (B.6)$$

  If $C_{\text{failed}}(\bar{x}) \cap C_{\text{CM}} = \emptyset$, $\bar{x}' = \bar{x}$, thus $\bar{x}' \rightarrow \bar{x}$ in this case. If $C_{\text{failed}}(\bar{x}) \cap C_{\text{CM}} \neq \emptyset$, we order this set according to the permutation $\sigma$ that satisfies assumption (A) of Theorem 5.2.6. Let $\hat{x}_0, \hat{x}_1, \ldots, \hat{x}_{\hat{n}}$ denote a sequence of states such that $\hat{x}_0 := \bar{x}'$, $\hat{x}_{\hat{n}} := \bar{x}$, and $\hat{x}_k$, $1 \leq k \leq \hat{n}$, is the state which is obtained from $\hat{x}_{k-1}$ when the $k$-th common mode failure (according to $\sigma$) occurs. The order of $\sigma$ assures that the corresponding common mode failures can occur, because for each of these common mode failures there is at least one corresponding component in $C_{\text{indep,line}}$ whose affected line segments are operating (namely, by assumption (A), neither have all redundant elements failed)

\(^5\) Clearly,

$$n \leq \sum_{c \in C_{\text{failed}}(\bar{x}) \cap C_{\text{indep}}} m_c .$$
independently nor have they failed due to another common mode failure). Thus we have

\[ \tilde{p}(\hat{x}_k \mid \hat{x}_{k-1}, \mu(\hat{x}_{k-1})) \] (4.19), (5.9) \[ > 0 \quad 1 \leq k \leq \hat{n} \, . \] (B.7)

Using (B.7), we conclude from (A.1) that \( \bar{x}' \rightarrow \bar{x} \).

Finally, with (B.4) and (B.6) we conclude from (A.2) that \( x^0 \rightarrow \bar{x} \).

(c) From (a) we have \( x^0 \rightarrow x^0 \) and from (b) we have \( x^0 \rightarrow \bar{x} \). Thus, we conclude from (A.2) that \( x^0 \rightarrow \bar{x} \).

\( \square \) (Lemma B.2.1)

**Lemma B.2.2** Let \( n_{\text{res}} \geq 1, \bar{x} \in S^{\text{transient}}, \) and \( \mu \in \Pi^{MD} \) a deterministic stationary policy. Then, for all \( N \geq 0 \) we have

\[ \mathbb{P}[\bar{X}_{\mu,N} = \bar{x} \mid \bar{X}_{\mu,0} = x^0, \mu] = 0 , \]

i.e., \( \bar{x} \) is not accessible from \( x^0 \) (for \( \bar{X}_{\mu} \)), where \( x^0 = (0, \ldots, 0) \).

**Proof of Lemma B.2.2:** As \( x^0 \in S^{\text{recurrent}} \), the claim is true for \( N = 0 \). In the following, let \( N \geq 1 \). We prove the lemma by contradiction. Assume that \( \bar{x} \) is accessible from \( x^0 \). Then, there exists an order of the independent and common mode failures (that are present in \( \bar{x} \)) in which they can occur with positive probability.

Let \( \sigma \) denote a permutation of \( \{1, \ldots, |C|\} \) such that \( c_{\sigma(1)}, \ldots, c_{\sigma(|C^{\text{CM}} \cap C^{\text{failed}}(\bar{x})|)} \) correspond to such an order of the common mode failures. As assumption (A) is violated for \( \bar{x} \), there exists \( j, 1 \leq j \leq |C^{\text{CM}} \cap C^{\text{failed}}(\bar{x})| \), such that

\[ C_{c_{\sigma(j)}}(\bar{x}) \setminus \bigcup_{k=1}^{j-1} C_{c_{\sigma(k)}}(\bar{x}) = \emptyset \, . \] (B.8)

For a state \( x \in S \), define

\[ E_{\text{line.on}}(x) := \bigcup_{c \in C^{\text{indepline}}} \bigcup_{\chi_c(x) = 1} E_{l_c} , \]
i.e., $E^{\text{line, on}}(x)$ contains all line segments of power supply lines that are operating in state $x$. Let $x' \in S$ denote the state that is attained immediately before the common mode failure corresponding to $c_{\sigma(j)}$ occurs. The common mode failure corresponding to $c_{\sigma(j)}$ can occur if and only if

$$E^{\text{line, on}}(x') \cap E_{c_{\sigma(j)}} \neq \emptyset.$$ 

First, we claim that

$$E^{\text{line, on}}(x') \cap E_{c_{\sigma(j)}} \subseteq \bigcup_{c \in C_{\sigma(j)}(x')} E_{l_c}. \quad (B.9)$$

Namely, any line segment $e$ belonging to the set on the left is affected by the common mode failure corresponding to $c_{\sigma(j)}$, i.e., this line segment belongs to $E_{l_c}$ of a component $c \in C_{\sigma(j)}$ (cf. (5.2) on page 108 and (5.3) on page 109). As $e$ belongs to an operating power supply line, we have $\chi_c(x') = 1$ and thus $x'_c \neq m_c$ (cf. Footnote 6 on page 117), i.e., $c \in C_{\sigma(j)}(x')$ and $e$ is an element of the set on the right.

After the occurrence of the common mode failure corresponding to $c_{\sigma(j)}$, no additional independent failures can occur in the components in $C_{c_{\sigma(j)}}$ (because none of the corresponding power supply lines is operating any more). As a consequence, we have $C_{c_{\sigma(j)}}(\hat{x}) = C_{c_{\sigma(j)}}(x')$ for any state $\hat{x}$ that is attained after $x'$. In particular, we have $C_{c_{\sigma(j)}}(\bar{x}) = C_{c_{\sigma(j)}}(x')$ and thus (from (B.9))

$$E^{\text{line, on}}(x') \cap E_{c_{\sigma(j)}} \subseteq \bigcup_{c \in C_{\sigma(j)}(x)} E_{l_c}. \quad (B.10)$$

If $j = 1$, we have $C_{c_{\sigma(1)}}(\bar{x}) = \emptyset$ by (B.8) and thus $E^{\text{line, on}}(x') \cap E_{c_{\sigma(1)}} = \emptyset$ from (B.10). Hence, the common mode failure corresponding to $c_{\sigma(1)}$ cannot occur, which is a contradiction. In the following, we assume that $j > 1$.

For $c \in C^{\text{CM}}$, we observe from the definition of $C_c(x)$ (cf. (5.14) on page 124) and from (5.2) (on page 108) that

$$E_c \supseteq \bigcup_{c' \in C_c(x)} E_{l_{c'}} \quad \forall x \in S. \quad (B.11)$$

As $E^{\text{line, on}}(x')$ cannot contain line segments that were affected by common mode failures corresponding to $c_{\sigma(1)}, \ldots, c_{\sigma(j-1)}$ (because these common
mode failures have already occurred in $x'$, we have

$$E^{\text{line, on}}(x') \subseteq E \setminus \bigcup_{k=1}^{j-1} E_{c_{\sigma(k)}}$$

**(B.11)**

$$\subseteq E \setminus \bigcup_{k=1}^{j-1} \bigcup_{c' \in C_{c_{\sigma(k)}}(\bar{x})} E_{l_{c'}}$$

$$= E \setminus \bigcup_{c' \in \bigcup_{k=1}^{j-1} C_{c_{\sigma(k)}}(\bar{x})} E_{l_{c'}}. \quad (B.12)$$

As $E_{c_{\sigma(j)}} \subseteq E$ we obtain from (B.12)

$$E^{\text{line, on}}(x') \cap E_{c_{\sigma(j)}} \subseteq E_{c_{\sigma(j)}} \setminus \bigcup_{c' \in \bigcup_{k=1}^{j-1} C_{c_{\sigma(k)}}(\bar{x})} E_{l_{c'}}.$$

Together with (B.10), we obtain

$$E^{\text{line, on}}(x') \cap E_{c_{\sigma(j)}} \subseteq \left( E_{c_{\sigma(j)}} \setminus \bigcup_{c' \in \bigcup_{k=1}^{j-1} C_{c_{\sigma(k)}}(\bar{x})} E_{l_{c'}} \right) \cap \bigcup_{c \in C_{c_{\sigma(j)}}(\bar{x})} E_{l_{c}}$$

$$= \left( E_{c_{\sigma(j)}} \cap \bigcup_{c \in C_{c_{\sigma(j)}}(\bar{x})} E_{l_{c}} \right) \setminus \bigcup_{c' \in \bigcup_{k=1}^{j-1} C_{c_{\sigma(k)}}(\bar{x})} E_{l_{c'}}$$

**(B.11)**

$$= \bigcup_{c \in C_{c_{\sigma(j)}}(\bar{x}) \setminus \bigcup_{k=1}^{j-1} C_{c_{\sigma(k)}}(\bar{x})} E_{l_{c}}.$$

By (B.8), the last union of components is empty and thus $E^{\text{line, on}}(x') \cap E_{c_{\sigma(j)}} = \emptyset$. Hence, the common mode failure corresponding to $c_{\sigma(j)}$ cannot occur, which is a contradiction.

□(Lemma B.2.2)

□
Theorem 5.3.1 (on page 136)

(i) For $0 \leq k \leq N$, $S^k \neq \emptyset$.

(ii) For $k > N$, $S^k = \emptyset$.

Proof of Theorem 5.3.1:

(i) Define the state $\hat{x} \in S$ component-wise by

$$\hat{x}_c := \begin{cases} m_c & \text{if } c \in C^{\text{indep}} \\ 0 & \text{otherwise.} \end{cases}$$

As $C^{\text{CM}} \cap C^{\text{failed}}(\hat{x}) = \emptyset$, $\hat{x} \in S^{\text{recurrent}}$ (cf. Theorem 5.2.6). Hence, as $\sum_{c \in C} \hat{x}_c = N$, $\hat{x} \in S^N$, i.e., $S^N \neq \emptyset$.

Now let $x' \in S$ be a state with $x'_c \leq \hat{x}_c$ for all $c \in C$. Then, $x'$ satisfies $C^{\text{CM}} \cap C^{\text{failed}}(x') = \emptyset$ and thus belongs to $S^{\text{recurrent}}$. By choosing the components’ states $x'_c$, $c \in C^{\text{indep}}$, (still with $x'_c \leq \hat{x}_c$) such that $\sum_{c \in C^{\text{indep}}} x'_c = k$ (which is possible for $0 \leq k \leq N$), the state $x'$ belongs to $S^k$, i.e., $S^k \neq \emptyset$.

(ii) Let $k > N$ and suppose there exists a state $x \in S^k$. Then we have $C^{\text{CM}} \cap C^{\text{failed}}(x) \neq \emptyset$ (because otherwise, $\sum_{c \in C} x_c \leq \sum_{c \in C^{\text{indep}}} m_c = N < k$, which contradicts $x \in S^k$). Let $\sigma$ denote a permutation of the set $\{1, \ldots, |C|\}$ that satisfies condition (A) of Theorem 5.2.6 (which exists as $x \in S^k \subseteq S^{\text{recurrent}}$). Condition (A) assures that

$$\left| C_{\sigma(j)}(x) \right| \geq 1 \quad 1 \leq j \leq |C^{\text{CM}} \cap C^{\text{failed}}(x)|, \quad \text{for all } 1 \leq j \leq |C^{\text{CM}} \cap C^{\text{failed}}(x)|. \quad \text{(Dj(x))}$$

Let $\beta(x) := |C^{\text{CM}} \cap C^{\text{failed}}(x)|$. As the sets $D_j(x)$, $1 \leq j \leq \beta(x)$, are disjoint subsets of the set $\{c \in C^{\text{indep}} \mid x_c \neq m_c\}$, we have

$$|\{c \in C^{\text{indep}} \mid x_c \neq m_c\}| \geq \sum_{j=1}^{\beta(x)} |D_j(x)| \geq \beta(x).$$
We thus obtain

\[
  k = \sum_{c \in C} x_c = \sum_{c \in C^{\text{CM}}} x_c + \sum_{c \in C^{\text{indep}}} x_c
\]

where \(C^{\text{CM}} = \{c \in C \mid \text{failed}(x) = \beta(x)\}\).

\[
  \leq \beta(x) + \sum_{c \in C^{\text{indep}}} m_c - \left\{ \sum_{c \in C^{\text{indep}}} \left[ x_c \neq m_c \right] \right\}
\]

\[
  \leq \sum_{c \in C^{\text{indep}}} m_c = N,
\]

which is a contradiction to \(k > N\) and we conclude that \(S^k = \emptyset\).

\[\blacksquare\]

**Theorem 5.3.3 (on page 138)** Let \(n_{\text{res}} \geq 1\) and assume (5.28), (5.29) hold. Then, the discrete-time MDP \(\tilde{M}_K, 0 \leq K \leq N\), is recurrent, i.e., for every deterministic stationary policy \(\mu \in \Pi^{\text{MD}}\) restricted to \(\bigcup_{k=0}^{K} S^k\), the embedded Markov chain \(\tilde{X}_{K,\mu}\) (of \(\tilde{M}_K\)) corresponding to \(\mu\) is irreducible and all states are recurrent.

**Proof of Theorem 5.3.3:** For \(K = N\), \(\tilde{M}_N\) coincides with \(\tilde{M}^{\text{rec}}\) (cf. Remark 5.3.2) which is recurrent (cf. (5.16)). Before we continue the proof, we make the following observation about \(\tilde{M}^{\text{rec}}\) (base model).

**Lemma B.2.3** Let \(n_{\text{res}} \geq 1\) and let \(\mu \in \Pi^{\text{MD}}\) (restricted to \(S^{\text{recurrent}}\)) be a deterministic stationary policy for \(\tilde{M}^{\text{rec}}\). Then for each \(k\) with \(0 \leq k \leq N - 1\) there exist \(x \in S^k\) and \(\bar{x} \in S^{k+1}\) such that \(\tilde{p}(\bar{x} \mid x, \mu(x)) > 0\).

**Proof of Lemma B.2.3:** By Theorem 5.3.1, \(S^k\) and \(S^{k+1}\) of the lemma are non-empty. As \(\tilde{M}^{\text{rec}}\) is recurrent (cf. (5.16)), \(\tilde{X}_\mu\), the embedded Markov chain of \(\tilde{M}^{\text{rec}}\) corresponding to \(\mu\), is irreducible. Thus, all states in \(S^{\text{recurrent}}\) communicate (for \(\tilde{X}_\mu\)) and the claim follows from (A.1) and (5.24). \(\blacksquare\)

We assume in the sequel that \(K < N\). Let \(\mu \in \Pi^{\text{MD}}\) be a deterministic stationary policy. From the proof of Lemma B.2.1 we conclude that for \(\tilde{X}_{K,\mu}\), each state in \(\bigcup_{k=0}^{K} S^k\) communicates with \(x^0\) and thus all states in \(\bigcup_{k=0}^{K} S^k\) communicate. Due to (5.28) and (5.29) we have \(s^1 \rightarrow x^0\) and \(s^k \leftrightarrow s^j\) for \(1 \leq k, j \leq N\) (cf. (A.1)). Due to Lemma B.2.3 (and as
there exist \( x \in S^K \) and \( \bar{x} \in S^{K+1} \) such that \( \tilde{p}(\bar{x} \mid x, \mu(x)) > 0 \). Hence \( \tilde{p}(s^{K+1} \mid x, \mu(x)) > 0 \) and \( x \to s^{K+1} \). We thus conclude (by (A.2)) that all states in \( \bigcup_{k=0}^K S^k \cup \{s^1, \ldots, s^N\} \) communicate (for \( \tilde{X}_{K, \mu} \)). As the state space is finite, all states are recurrent (cf. Theorem 4.1.1). □

---

Theorem 5.3.4 (on page 140) Let \( n_{\text{res}} \geq 1 \) and \( 0 \leq K \leq N \). Assume that the transition probabilities \( \tilde{p}(s^l \mid s^k) \) in the aggregated states satisfy (5.27), (5.28), (5.29), and

\[
\tilde{p}(s^{k+1} \mid s^k) \geq \max_{x \in S^k, u \in U(x)} \sum_{\bar{x} \in S^{k+1}} \tilde{p}(\bar{x} \mid x, u) \quad 1 \leq k \leq N - 1 \quad (B.13)
\]

\[
\tilde{p}(s^{k-1} \mid s^k) \leq \min_{x \in S^k, u \in U(x)} \sum_{\bar{x} \in S^{k-1}} \tilde{p}(\bar{x} \mid x, u) \quad 1 \leq k \leq N \quad . (B.14)
\]

Further assume that the cost of the aggregated states satisfy

\[
\tilde{g}(s^k) \geq \max_{x \in S^k, u \in U(x)} \tilde{g}(x, u) \quad 1 \leq k \leq N ,
\]

and

\[
\tilde{g}(s^{k+1}) \geq \tilde{g}(s^k) \geq \tilde{g}(x^0, u^0) \quad 1 \leq k \leq N - 1 .
\]

Let \( \mu \in \Pi^{\text{MD}} \) be a deterministic stationary policy. Then the average expected cost \( \lambda_{\tilde{M}_K, \mu} \) of \( \tilde{M}_K \) gives an upper bound on the average expected cost \( \lambda_{\tilde{M}_{\text{rec}}, \mu} \) (and \( \lambda_{M, \mu} \)) of \( \mu \) of \( \tilde{M}_{\text{rec}} \) (and \( M \)), i.e.,

\[
\lambda_{M, \mu} \overset{(5.17)}{=} \lambda_{\tilde{M}_{\text{rec}}, \mu} \leq \lambda_{\tilde{M}_K, \mu} .
\]

In addition, the optimal average expected cost \( \lambda^*_{\tilde{M}_K} \) of \( \tilde{M}_K \) gives an upper bound on the optimal average expected cost of \( \tilde{M}_{\text{rec}} \) (and thus also of the base model \( M \)), i.e.,

\[
\lambda^*_{M} \overset{(5.18)}{=} \lambda^*_{\tilde{M}_{\text{rec}}} \leq \lambda^*_{\tilde{M}_K} .
\]

---

Proof of Theorem 5.3.4 (details; \( K = 0 \)):

Before we start with the proof of \( K = 0 \), we make the following observation about \( \tilde{M}_{\text{rec}} \) (base model).

\[6\] For \( \tilde{M}_K, \mu \) is restricted to \( \bigcup_{k=0}^K S^k \) and for \( \tilde{M}_{\text{rec}}, \mu \) is restricted to \( S_{\text{current}}^{\text{recurrent}} \).
Lemma B.2.4 Let $n_{res} \geq 1$ and let $\mu \in \Pi^{MD}$ (restricted to $S^{\text{recurrent}}$) be a deterministic stationary policy for $\mathcal{M}^{\text{rec}}$. Then for each $k$ with $1 \leq k \leq N$ and for all $x \in S^k$, there exists $\bar{x} \in S^{k-1}$ such that $\tilde{p}(\bar{x} \mid x, \mu(x)) > 0$.

Proof of Lemma B.2.4: The claim follows from the argument leading to (B.3) in the proof of Lemma B.2.1 by observing that $x_0, x_1$ in (B.3) satisfy: $x_0 \in S^k \Rightarrow x_1 \in S^{k-1}$ for $1 \leq k \leq N$ (due to (5.24) and because $S^{\text{recurrent}}$ is closed for $\tilde{X}_\mu$). □ (Lemma B.2.4)

In the sequel, we assume that $K = 0$. The idea of the proof of the theorem is given on page 141. We also use some of the notations that were introduced there. The following proof is based on [Zen10].

Let $\mu \in \Pi^{MD}$ be a deterministic stationary policy. For $K = 0$, all states are aggregated and there is no choice of control in the aggregated model (cf. Remark 5.3.5), i.e., $\mu$ has no influence in the aggregated model. We thus write $\tilde{X}_K$ instead of $\tilde{X}_{K,\mu}$ for the embedded Markov chain of $\tilde{M}_K$ with transition probabilities $\tilde{p}(s^l \mid s^k)$ in aggregated states. $\tilde{X}_K$ is a discrete-time homogeneous Markov chain on the set of states $\{s^0, s^1, \ldots, s^N\}$.

“Exact aggregation” (parts (b), (c) of the proof):

Recall that $\nu_\mu$ denotes the stationary distribution of $\tilde{X}_\mu$, the embedded Markov chain of $\tilde{M}^{\text{rec}}$ (base model) corresponding to $\mu$ (where $\mu$ is restricted to $S^{\text{recurrent}}$). We have $\nu_\mu(x) \neq 0$ for all $x \in S^{\text{recurrent}}$ (cf. (a) of the idea of the proof, page 141). Define the transition probabilities $\tilde{p}^{\text{exact}}_\mu(s^l \mid s^k)$ by

$$\tilde{p}^{\text{exact}}_\mu(s^l \mid s^k) := \frac{\sum_{x \in S^k} \nu_\mu(x) \sum_{\bar{x} \in S^l} \tilde{p}(\bar{x} \mid x, \mu(x))}{\sum_{x \in S^k} \nu_\mu(x)} \quad 0 \leq k, l \leq N,$$

and let $\nu^{\text{exact}}_\mu$ denote the stationary distribution of the resulting discrete-time homogeneous Markov chain (denoted by $\tilde{X}^{\text{exact}}_\mu$) on the set of states $\{s^0, s^1, \ldots, s^N\}$. From Theorem A.1.5 we have

$$\nu^{\text{exact}}_\mu(s^k) = \sum_{x \in S^k} \nu_\mu(x) \quad 0 \leq k \leq N. \quad (B.15)$$

Observe that $\tilde{p}^{\text{exact}}_\mu(s^l \mid s^k)$ satisfy (5.27) (due to (5.24)) as well as (5.28) and (5.29) (due to Lemma B.2.3 (from the proof of Theorem 5.3.3) and Lemma B.2.4, respectively). Further, $\tilde{p}^{\text{exact}}_\mu(s^1 \mid s^0) = \sum_{\bar{x} \in S^1} \tilde{p}(\bar{x} \mid \bar{x} \in S^1).$
\( x^0, \mu(x^0) \), i.e., (5.26) holds (cf. (5.30)). Consequently, \( \tilde{p}^\text{exact}_\mu(s^l | s^k) \) satisfy the requirements on the transition probabilities of \( \tilde{\mathcal{M}}_K \). Hence, \( \tilde{\mathcal{X}}^\text{exact}_K \) is the embedded Markov chain in aggregated states of \( \tilde{\mathcal{M}}_K \) with transition probabilities \( \tilde{p}^\text{exact}_\mu(s^l | s^k) \) and as (B.15) establishes (5.37), part (b) is finished.

Recall that \( \lambda^\text{exact}_{\tilde{\mathcal{M}}_K, \mu} \) denotes the (optimal\(^7\)) average expected cost of \( \tilde{\mathcal{M}}_K \) with transition probabilities \( \tilde{p}^\text{exact}_\mu(s^l | s^k) \) and that \( \lambda^\text{exact}_{\tilde{\mathcal{M}}_K, \mu} \) denotes the average expected cost of \( \mu \) for \( \tilde{\mathcal{M}}_K \). For notational convenience, define

\[
\tilde{g}(s^0) := \tilde{g}(x^0, u^0),
\]
i.e., \( \tilde{g}(s^0) = \tilde{g}(x^0, \hat{\mu}(x^0)) \) for any deterministic stationary policy \( \hat{\mu} \in \Pi^{\text{MD}} \) (cf. (5.30)). We have

\[
\lambda^\text{exact}_{\tilde{\mathcal{M}}_K, \mu} = \sum_{x \in S^{\text{recurrent}}} \nu_\mu(x) \cdot \tilde{g}(x, \mu(x)) = \sum_{k=0}^{N} \sum_{x \in S^k} \nu_\mu(x) \cdot \tilde{g}(x, \mu(x)) \leq \tilde{g}(s^k) \quad \forall x \in S^k \leq \sum_{k=0}^{N} \nu_\mu^\text{exact}(s^k) \cdot \tilde{g}(s^k) = \lambda^\text{exact}_{\tilde{\mathcal{M}}_K, \mu},
\]

which finishes part (c).

“Modification” (part (e) of the proof):

Assume that the transition probabilities \( \tilde{p}(s^l | s^k) \) of \( \tilde{\mathcal{M}}_K \) satisfy the assumptions in Theorem 5.3.4. For notational convenience, we define

\[
\tilde{p}(s^1 | s^0) := \tilde{p}(s^1 | x^0, u^0) \overset{(5.26)}{=} \sum_{x \in S^1} \tilde{p}(x | x^0, u^0),
\]
i.e., \( \tilde{p}(s^1 | s^0) = \tilde{p}(s^1 | x^0, \hat{\mu}(x^0)) \) for any deterministic stationary policy \( \hat{\mu} \in \Pi^{\text{MD}} \) (cf. (5.30)). From (B.13) and (B.14) and the definition of \( \tilde{p}^\text{exact}_\mu(s^l | s^k) \) we observe

\[
\tilde{p}^\text{exact}_\mu(s^{k+1} | s^k) \leq \tilde{p}(s^{k+1} | s^k) \quad 0 \leq k \leq N - 1
\]
\[
\tilde{p}^\text{exact}_\mu(s^{k-1} | s^k) \geq \tilde{p}(s^{k-1} | s^k) \quad 1 \leq k \leq N.
\]

\(^7\) because for \( K = 0 \), there is no choice of control.
Recall that $\tilde{X}_K$ denotes the embedded Markov chain of $\tilde{M}_K$ with transition probabilities $\tilde{p}(s^l \mid s^k)$ and let $\nu_K$ denote the stationary distribution of $\tilde{X}_K$.

**Lemma B.2.5**  
(i) The stationary distribution $\nu_K$ of $\tilde{X}_K$ is given by

$$
\nu_K(s^0) = \frac{1}{1 + \sum_{k=1}^{N} \prod_{j=0}^{k-1} \tilde{p}(s^{j+1} \mid s^j)},
$$

$$
\nu_K(s^k) = \prod_{j=0}^{k-1} \frac{\tilde{p}(s^{j+1} \mid s^j)}{\tilde{p}(s^j \mid s^{j+1})} \cdot \nu_K(s^0), \quad 1 \leq k \leq N.
$$

and analogously for $\nu^\mu_{\text{exact}}$ (the stationary distribution of $\tilde{X}^\mu_{\text{exact}}$) with $\tilde{p}^\mu_{\text{exact}}(s^l \mid s^k)$ instead of $\tilde{p}(s^l \mid s^k)$.

(ii) For $0 \leq k \leq N - 1$ we have

$$
\nu^\mu_{\text{exact}}(s^k) \leq \nu_K(s^k) \Rightarrow \nu^\mu_{\text{exact}}(s^{k+1}) \leq \nu_K(s^{k+1}).
$$

**Proof of Lemma B.2.5:**

(i) Let $P$ denote the transition matrix of $\tilde{X}_K$. From the necessary condition $\nu^T_K = \nu^T_K P$ for a stationary distribution, we obtain by induction for $N - 1 \geq k \geq 0$ (because $P$ is a tridiagonal matrix):

$$
\nu_K(s^{k+1}) = \frac{\tilde{p}(s^{k+1} \mid s^k)}{\tilde{p}(s^k \mid s^{k+1})} \nu_K(s^k), \quad (B.16)
$$

which implies the second equation of (i). The first equation follows with the condition $\sum_{k=0}^{N} \nu_K(s^k) = 1$. As the transition matrix of $\tilde{X}^\mu_{\text{exact}}$ has exactly the same structure as $P$, the formulas also hold for $\nu^\mu_{\text{exact}}$ (with $\tilde{p}^\mu_{\text{exact}}(s^l \mid s^k)$ instead of $\tilde{p}(s^l \mid s^k)$).

(ii) Consider (using (B.16), adapted to $\nu^\mu_{\text{exact}}$ and $\tilde{p}^\mu_{\text{exact}}(s^l \mid s^k)$):

$$
\nu^\mu_{\text{exact}}(s^{k+1}) = \frac{\tilde{p}^\mu_{\text{exact}}(s^{k+1} \mid s^k)}{\tilde{p}^\mu_{\text{exact}}(s^k \mid s^{k+1})} \cdot \nu^\mu_{\text{exact}}(s^k)
$$

which implies

$$
\nu^\mu_{\text{exact}}(s^{k+1}) \leq \nu_K(s^{k+1}).
$$

\hfill (Lemma B.2.5)
Recall that $\lambda_{\tilde{M}_K,\mu}^{\text{exact}}$ denotes the (optimal) average expected cost of $\tilde{M}_K$ with transition probabilities $\tilde{p}^{\text{exact}}_\mu(s^l \mid s^k)$.

We write $\lambda_{\tilde{M}_K}^{\text{exact}}$ (instead of $\lambda_{\tilde{M}_K,\mu}^{\text{exact}}$) for the (optimal) average expected cost of $\tilde{M}_K$ (with transition probabilities $\tilde{p}(s^l \mid s^k)$). We have

$$
\lambda_{\tilde{M}_K,\mu}^{\text{exact}} - \lambda_{\tilde{M}_K}^{\text{exact}} = \sum_{k=0}^N (\nu^{\text{exact}}_\mu(s^k) - \nu(s^k))\tilde{g}(s^k).
$$

Case 1: $\nu^{\text{exact}}_\mu(s^k) - \nu(s^k) \geq 0$ for all $0 \leq k \leq N$. Assume there exists $\hat{k}, 0 \leq \hat{k} \leq N$, such that $\nu^{\text{exact}}_\mu(s^{\hat{k}}) - \nu(s^{\hat{k}}) > 0$. Then we have

$$
1 = \sum_{\substack{k=0 \\ k \neq \hat{k}}}^N \nu^{\text{exact}}_\mu(s^k) + \nu^{\text{exact}}_\mu(s^{\hat{k}}) > \sum_{k=0}^N \nu(s^k) = 1,
$$

which is a contradiction. Hence $\nu^{\text{exact}}_\mu(s^k) = \nu(s^k)$ for all $0 \leq k \leq N$. From (B.17) we thus conclude

$$
\lambda_{\tilde{M}_K,\mu}^{\text{exact}} - \lambda_{\tilde{M}_K}^{\text{exact}} = 0.
$$

Case 2: There exists $0 \leq k \leq N$ such that $\nu^{\text{exact}}_\mu(s^k) - \nu(s^k) < 0$. Assume $\nu^{\text{exact}}_\mu(s^0) - \nu(s^0) < 0$. By Lemma B.2.5(ii), we have $\nu^{\text{exact}}_\mu(s^k) - \nu(s^k) < 0$ for all $0 \leq k \leq N$ and thus

$$
1 = \sum_{k=0}^N \nu^{\text{exact}}_\mu(s^k) < \sum_{k=0}^N \nu(s^k) = 1,
$$

which is a contradiction. Thus $\nu^{\text{exact}}_\mu(s^0) - \nu(s^0) \geq 0$. Define

$$
k_0 := \min\{k \mid \nu^{\text{exact}}_\mu(s^k) - \nu(s^k) < 0\}.
$$

By the previous argument, $k_0 \geq 1$. Hence $\nu^{\text{exact}}_\mu(s^k) - \nu(s^k) \geq 0$ for all $0 \leq k < k_0$ and (by Lemma B.2.5(ii)) $\nu^{\text{exact}}_\mu(s^k) - \nu(s^k) < 0$ for
$k_0 \leq k \leq N$. We obtain

$$
\lambda_{\mathcal{M}_{K,\mu}} - \lambda_{\tilde{\mathcal{M}}_K} \overset{(B.17)}{=} \sum_{0 \leq k < k_0} (\nu_{\mu}^{\text{exact}}(s^k) - \nu_K(s^k)) \tilde{g}(s^k) \geq 0
$$

$$
+ \sum_{k_0 \leq k \leq N} (\nu_{\mu}^{\text{exact}}(s^k) - \nu_K(s^k)) \tilde{g}(s^k) < 0
$$

$$
\leq \tilde{g}(s^{k_0}) \sum_{0 \leq k \leq N} (\nu_{\mu}^{\text{exact}}(s^k) - \nu_K(s^k)) \geq \tilde{g}(s^{k_0})
$$

$$
= \tilde{g}(s^{k_0})(\sum_{0 \leq k \leq N} \nu_{\mu}^{\text{exact}}(s^k) - \sum_{0 \leq k \leq N} \nu_K(s^k)) = 0.
$$

Thus, we proved that $\lambda_{\mathcal{M}_{K,\mu}}^{\text{exact}} \leq \lambda_{\tilde{\mathcal{M}}_K}$, which finishes part (e).

Proof of Theorem 5.3.4 (details; $K > 0$): We assume that $1 \leq K \leq N - 1$ (and that $N \geq 2$). The idea of the proof is given on page 141. In the following, the key properties are marked with a frame.

Let $\mu \in \Pi^{\text{MD}}$ be a deterministic stationary policy. Recall that $\tilde{\mathcal{X}}_{K,\mu}$ denotes the embedded Markov chain of $\tilde{\mathcal{M}}_K$ corresponding to $\mu$ (where $\mu$ is restricted to $\bigcup_{k=0}^{K} S^k$). For the ease of presentation, we omit the dependency on $\mu$ in some of the notations.

Notations and first properties

Figure B.2 shows the structure of the transition matrix $P_{K,\mu}$ of $\tilde{\mathcal{X}}_{K,\mu}$. The entries with a "*" are the (a priori arbitrary) transition probabilities $\tilde{p}(s^I \mid s^k)$ in the aggregated states. We write $p_{k,i}$ for $\tilde{p}(s^I \mid s^k)$. The figure
also illustrates the definitions of the following quantities.

\[ D := \bigcup_{k=1}^{K} S^k \]

\[ p_{0,0} := \tilde{p}(x^0 \mid x^0, \mu(x^0)) \in \mathbb{R} \]

\[ p_{0,\cdot} \in \mathbb{R}^{1 \times |D|} \text{ with } (p_{0,\cdot})_x := \tilde{p}(x \mid x^0, \mu(x^0)) \text{ for } x \in D \]

\[ p_{\cdot,0} \in \mathbb{R}^{|D| \times 1} \text{ with } (p_{\cdot,0})_x := \tilde{p}(x^0 \mid x, \mu(x)) \text{ for } x \in D \]

\[ p_{\cdot,s_{K+1}} \in \mathbb{R}^{|S^K| \times 1} \text{ with } (p_{\cdot,s_{K+1}})_x := \tilde{p}(s_{K+1} \mid x, \mu(x)) \text{ for } x \in S^K \]

\[ P_D \in \mathbb{R}^{|D| \times |D|} \text{ with } (P_D)_{x,\bar{x}} := \tilde{p}(\bar{x} \mid x, \mu(x)) \text{ for } x, \bar{x} \in D. \]

\[ \text{cf. (5.26).} \]
For any stationary distribution $\zeta$ that is defined on the states $x \in D$, let $\zeta_s^k$, $1 \leq k \leq K$, denote the (column) vector of the stationary probabilities of all states in $S^k$ and let

$$
\zeta_D := \begin{pmatrix}
\zeta^1_s \\
\vdots \\
\zeta^K_s
\end{pmatrix}.
$$

(B.18)

The vector $e := (1, \ldots, 1)^T$ denotes the (column) vector of 1’s and $I$ the identity matrix (in the “appropriate” dimension).

Let $\nu$ denote the stationary distribution of $\tilde{X}_{K,\mu}$ with transition probabilities $p_{k,l}$ in the aggregated states. The condition $\nu^T = \nu^T P_{K,\mu}$ on the stationary distribution yields

$$
\nu_D = \nu(x^0)p_0^T + P_D^T \nu_D,
$$

i.e.,

$$(I - P_D^T)\nu_D = \nu(x^0)p_0^T .
$$

**Lemma B.2.6** The inverse $(I - P_D^T)^{-1}$ exists and $(I - P_D^T)^{-1} \geq 0$.

The proof of Lemma B.2.6 is provided below. By Lemma B.2.6 we have

$$
\nu_D = \nu(x^0)(I - P_D^T)^{-1} p_0^T .
$$

(B.19)

**Proof of Lemma B.2.6:** The proof relies on properties of $M$-matrices, cf. [BP79, Definition 6.1.2]. In the following proof, references to definitions and theorems that are marked with $(\star)$ all refer to the book [BP79]. By Theorem 8.4.2$(\star)$, $I - P_{K,\mu}$ is an M-matrix (recall that $P_{K,\mu}$ is the transition matrix of the DTHMC $\tilde{X}_{K,\mu}$).

As $\tilde{M}_K$ is recurrent (cf. Theorem 5.3.3), $\tilde{X}_{K,\mu}$ is an irreducible DTHMC. Thus, by Theorem 8.3.9$(\star)$, $P_{K,\mu}$ is an irreducible matrix (cf. Definition 2.1.2$(\star)$). By Theorem 2.2.7$(\star)$, a square matrix $A \in \mathbb{R}^{n \times n}$ is irreducible if and only if the associated directed graph $G(A)$ (with a directed edge from node $i$ to node $j$ if and only if $A_{ij} \neq 0$, cf. Definition 2.2.4$(\star)$) is strongly connected (i.e., for any two nodes $i, j$ of $G(A)$ there exists a sequence of directed edges that leads from $i$ to $j$, cf. Definition 2.2.6$(\star)$).

As $G(P_{K,\mu})$ is strongly connected and as $P_{K,\mu}$ is a stochastic matrix
in $\mathbb{R}^{n \times n}$ with $n > 1$, each diagonal element $(P_{K,\mu})_{i,i}$ of $P_{K,\mu}$ satisfies $(P_{K,\mu})_{i,j} < 1$. Consequently, $(P_{K,\mu})_{i,j} \neq 0 \Rightarrow (I - P_{K,\mu})_{i,j} \neq 0$ for all $i, j$, i.e., $G(I - P_{K,\mu})$ contains all edges of $G(P_{K,\mu})$. Thus $G(I - P_{K,\mu})$ is also strongly connected and the matrix $I - P_{K,\mu}$ is irreducible. Further, as $(I - P_{K,\mu})e = 0$, $I - P_{K,\mu}$ is singular. Summing up, $I - P_{K,\mu}$ is a singular irreducible M-matrix. By Theorem 6.4.16(*), $I - P_D$ (which is a principal submatrix of $I - P_{K,\mu}$ with $I - P_D \neq I - P_{K,\mu}$) is a non-singular M-matrix and thus, by condition $N_{38}$ of Theorem 6.2.3(*), $(I - P_D)^{-1}$ exists and $(I - P_D)^{-1} \geq 0$. As $((I - P_D)^{-1})^T = ((I - P_D)^T)^{-1} = (I - P_D^T)^{-1}$, the lemma is proven.

□(Lemma B.2.6)

“Exact aggregation” (parts (b), (c) of the proof):

Recall that $\tilde{X}_\mu$ denotes the embedded Markov chain of $\tilde{M}_{\text{rec}}$ corresponding to $\mu$ (where $\mu$ is restricted to $S^{\text{recurrent}}$). We introduce some more notation.

$$A := \bigcup_{k=K+1}^{N} S^k = S \setminus (\{x^0\} \cup D)$$

$$P_{DA} \in \mathbb{R}^{|D| \times |A|} \text{ with } (P_{DA})_{x,\bar{x}} := \tilde{p}(\bar{x} \mid x, \mu(x)) \text{ for } x \in D, \bar{x} \in A$$

$$P_{AD} \in \mathbb{R}^{|A| \times |D|} \text{ with } (P_{AD})_{x,\bar{x}} := \tilde{p}(\bar{x} \mid x, \mu(x)) \text{ for } x \in A, \bar{x} \in D$$

$$P_A \in \mathbb{R}^{|A| \times |A|} \text{ with } (P_A)_{x,\bar{x}} := \tilde{p}(\bar{x} \mid x, \mu(x)) \text{ for } x, \bar{x} \in A.$$  

Using these as well as the previously defined notations, the transition matrix $P_\mu$ of $\tilde{X}_\mu$ is given by

$$P_\mu = \begin{pmatrix}
  x^0 & D & A \\
  D & p_{0,0} & p_{0,} & 0 \\
  A & p_{0,} & P_D & P_{DA} \\
  0 & P_{DA} & P_A
\end{pmatrix} \quad \text{(B.20)}$$

where the notations outside the box are used to indicate the corresponding states. The 0’s in $P_\mu$ are due to (5.27), using that $D \neq \emptyset$ because $K \geq 1$.

The following idea of duplicating states and the subsequent aggregation is taken from [MdSeSG89]. Based on $\tilde{X}_\mu$, define the following auxiliary discrete-time homogeneous Markov chain (cf. Figure B.3(a), (b)): duplicate all states in $D = \bigcup_{k=1}^{K} S^k$. The set of duplicated states is de-
noted by $D'$. The state $x' \in D'$ denotes the duplicate of $x \in D$. The transition matrix $P_{p}^{aux} = (p_{aux}(\bar{x} \mid x))_{x,\bar{x} \in S \cup D'}$ of the auxiliary DTHMC (on the set of states $S \cup D'$) is defined as follows:

$$
P_{p}^{aux} :=
\begin{array}{cccc}
x^0 & D & A & D' \\
p_{0,0} & p_{0,} & 0 & 0 \\
p_{0,0} & 0 & P_{DA} & 0 \\
p_{0,} & 0 & P_{D} & P_{DA} \\
p_{0,} & P_{DA} & P_{D} & 0 \\
\end{array}
$$

where the notations outside the box are again used to indicate the corresponding states. Note that $P_{p}^{aux}$ is indeed a stochastic matrix.
Let $\xi$ denote the stationary distribution of $\hat{X}_\mu$ and let $\nu^{aux}$ denote the stationary distribution of the auxiliary DTHMC.\(^9\)

**Lemma B.2.7** The stationary distributions $\xi$ and $\nu^{aux}$ satisfy:

\[
\nu^{aux}(x) = \xi(x) \quad x \in \{x^0\} \cup A
\]

\[
\nu^{aux}(x) + \nu^{aux}(x') = \xi(x) \quad x \in D.
\]

**Proof of Lemma B.2.7:** For the ease of notation, we write $\nu$ instead of $\nu^{aux}$. Similar to (B.18), let $\xi_A$, $\nu_A$, and $\nu_D'$ denote the column vectors of the stationary probabilities of states in $A$ and $D'$, respectively. From the condition $\nu^T = \nu^{aux} P_{aux}^T \mu$ we obtain

\[
\nu(x^0) = \nu(x^0)p_{0,0} + \nu_D p_{0,0} + \nu_D' p_{0,0}
\]

\[
\nu_D^T = \nu(x^0)p_{0,0} + (\nu_D + \nu_D')^T p_{0,0} \quad \text{(B.21)}
\]

\[
\nu_A^T = \nu_D^T P_{DA} + \nu_A^T P_A + \nu_D' P_{DA}
\]

\[
\nu_D'^T = \nu_A^T P_{AD} + \nu_D'^T P_D. \quad \text{(B.24)}
\]

Adding (B.22) and (B.24) yields

\[
(\nu_D + \nu_D')^T = \nu(x^0)p_{0,0} + (\nu_D + \nu_D')^T P_D + \nu_A^T P_{DA}. \quad \text{(B.25)}
\]

Define $\hat{\xi} \in \mathbb{R}^{\left|S\right|}$ by

\[
\hat{\xi}(x^0) := \nu(x^0)
\]

\[
\hat{\xi}_D := \nu_D + \nu_D'
\]

\[
\hat{\xi}_A := \nu_A.
\]

Then, due to (B.21), (B.25), and (B.23), $\hat{\xi}$ satisfies $\hat{\xi}^T = \hat{\xi}^T P_\mu$ and thus $\xi = \hat{\xi}$. \(\square\) (Lemma B.2.7)

Now the duplicated states in $D'$ as well as the states in $A = \bigcup_{k=K+1}^{N} S^k$ are aggregated as follows. For $1 \leq k \leq K$, the states $\{x' \in D' \mid x \in S^k\}$

---

\(^9\) For the auxiliary DTHMC, $x^0$ is accessible from each $x \in S \cup D'$ (cf. Lemma B.2.4 on page 201). Thus, $x^0$ is recurrent and there exists only one recurrent communicating class (which is closed), i.e., the auxiliary DTHMC is unichain. Thus, $\nu^{aux}$ is unique (cf. Theorem 4.1.4). Note, however, that some states $x \in D'$ may be transient for the auxiliary DTHMC.
are aggregated into $s^k$ and for $K + 1 \leq k \leq N$, the states $x \in S^k$ are aggregated into $s^k$ (cf. Figure B.3(c)). Define the transition probabilities between aggregated states as follows ($0 \leq l \leq N$):$^{10}$

$$
\begin{align*}
\tilde{p}_\mu^{\text{exact}}(s^l | s^k) &:= \frac{\sum_{x \in S^k} \nu^{\text{aux}}(x) \sum_{\bar{x} \in S^l} \tilde{p}(\bar{x} | x, \mu(x))}{\sum_{x \in S^k} L^{\text{aux}}(x)} \quad 1 \leq k \leq K \\
\tilde{p}_\mu^{\text{exact}}(s^l | s^k) &:= \frac{\sum_{x \in S^k} \nu^{\text{aux}}(x) \sum_{\bar{x} \in S^l} \tilde{p}(\bar{x} | x, \mu(x))}{\sum_{x \in S^k} L^{\text{aux}}(x)} \quad K + 1 \leq k \leq N.
\end{align*}
$$

The transition probability from $x \in S^K$ to $S^{K+1}$ is defined by (5.26) and the transition probabilities between non-aggregated states $x, \bar{x} \in \{x^0\} \cup D$ are given by $\tilde{p}(\bar{x} | x, \mu(x))$ (or, equivalently, by $\tilde{p}_\mu^{\text{aux}}(\bar{x} | x)$).

Due to (5.24), $\tilde{p}_\mu^{\text{exact}}(s^l | s^k)$ satisfy (5.27). Further, $\tilde{p}_\mu^{\text{exact}}(s^l | s^k)$ satisfy (5.28) and (5.29) due to Footnote 10, Lemma B.2.3 from the proof of Theorem 5.3.3, and Lemma B.2.4 on page 201. Consequently, $\tilde{p}_\mu^{\text{exact}}(s^l | s^k)$ satisfy the requirements on the transition probabilities in aggregated states of $\tilde{M}_K$. Hence the constructed aggregated DTHMC (which we denote by $\tilde{X}^{\text{exact}}_\mu$) corresponds to the embedded Markov chain $\tilde{X}^{\text{exact}}_{K,\mu}$ of $\tilde{M}_K$ with transition probabilities $\tilde{p}_\mu^{\text{exact}}(s^l | s^k)$ in aggregated states.

Let $\nu^{\text{exact}}$ denote the stationary distribution of $\tilde{X}^{\text{exact}}_\mu$. The following lemma establishes (5.37) and thus finishes part (b).

---

$^{10}$ For the auxiliary DTHMC, each $x \in S$ is accessible from $x^0$ (because $\tilde{X}_\mu$ is irreducible) and thus each $x \in S$ communicates with the recurrent state $x^0$ (cf. Footnote 9). Hence, each $x \in S$ is recurrent for the auxiliary DTHMC and $\nu^{\text{aux}}(x) > 0$ for all $x \in S$ (because the auxiliary DTHMC is unichain). Let $x_{(K+1)} \in S^{K+1}$ ($S^{K+1} \neq \emptyset$ due to Theorem 5.3.1). Due to Lemma B.2.4 on page 201, there exist states $x_{(K)} \in S^K, x_{(K-1)} \in S^{K-1}, \ldots, x_{(1)} \in S^1$ such that $\tilde{p}(x_{(k)} | x_{(k+1)}, \mu(x_{(k+1)})) > 0$ for $1 \leq k \leq K$ and $\tilde{p}(x^0 | x_{(1)}, \mu(x_{(1)})) > 0$. As a consequence, the corresponding duplicated states $x'_{(K)}, x'_{(K-1)}, \ldots, x'_{(1)} \in D'$ communicate with the recurrent state $x^0$ and thus $x'_{(k)}, 1 \leq k \leq K$, are recurrent for the auxiliary DTHMC. Hence, $\nu^{\text{aux}}(x'_{(k)}) > 0$ for $1 \leq k \leq K$ and thus $\sum_{x \in S^k} \nu^{\text{aux}}(x') > 0$ for $1 \leq k \leq K$. We further note that the states $x_{(K)}, x_{(K-1)}, \ldots, x_{(1)}$ can be chosen such that $\tilde{p}(x_{(k+1)} | x_{(k)}, \mu(x_{(k)})) > 0$ for $1 \leq k \leq K$ (the repaired components can fail again; cf. condition (A) of Theorem 5.2.6).
Lemma B.2.8 The stationary distribution $\nu^{\text{exact}}$ satisfies:

$$
\begin{align*}
\nu^{\text{exact}}(x^0) &= \xi(x^0) \\
\nu^{\text{exact}}(s^k) + \sum_{x \in S^k} \nu^{\text{exact}}(x) &= \sum_{x \in S^k} \xi(x) \quad 1 \leq k \leq K \\
\nu^{\text{exact}}(s^k) &= \sum_{x \in S^k} \xi(x) \quad K + 1 \leq k \leq N.
\end{align*}
$$

Proof of Lemma B.2.8: We first note that the transition probabilities of the constructed aggregated DTHMC $\tilde{X}^{\text{exact}}_\mu$ indeed satisfy the “aggregation formula” (A.7) of Theorem A.1.5. To see this, observe that for a duplicated state $x' \in D'$ with $x \in S^k$, $1 \leq k \leq K$, we have

$$
\sum_{\bar{x} \in S^l} \tilde{p}^{\text{aux}}(\bar{x} \mid x, \mu(x)) = \begin{cases} 
\tilde{p}^{\text{aux}}_{\mu}(x^0 \mid x') & l = 0 \\
\sum_{\bar{x} \in S^l} \tilde{p}^{\text{aux}}_{\mu}(\bar{x} \mid x') & 1 \leq l \leq K \\
\sum_{\bar{x} \in S^l} \tilde{p}^{\text{aux}}_{\mu}(\bar{x} \mid x') & K + 1 \leq l \leq N,
\end{cases}
$$

and for $x \in S^k$, $K + 1 \leq k \leq N$, we have

$$
\sum_{\bar{x} \in S^l} \tilde{p}(\bar{x} \mid x, \mu(x)) = \begin{cases} 
\tilde{p}^{\text{aux}}_{\mu}(x^0 \mid x) (= 0) & l = 0 \\
\sum_{\bar{x} \in S^l} \tilde{p}^{\text{aux}}_{\mu}(\bar{x} \mid x) & 1 \leq l \leq K \\
\sum_{\bar{x} \in S^l} \tilde{p}^{\text{aux}}_{\mu}(\bar{x} \mid x) & K + 1 \leq l \leq N.
\end{cases}
$$

Finally, $\sum_{\bar{x} \in S^{K+1}} \tilde{p}(\bar{x} \mid x, \mu(x)) = \sum_{\bar{x} \in S^{K+1}} \tilde{p}^{\text{aux}}_{\mu}(\bar{x} \mid x)$ for all $x \in S^K$.

From Theorem A.1.5 we thus obtain (the states which are not aggregated are considered as macrostates containing exactly one microstate)\(^{11}\)

$$
\begin{align*}
\nu^{\text{exact}}(x^0) &= \nu^{\text{aux}}(x^0) \\
\nu^{\text{exact}}(x) &= \nu^{\text{aux}}(x) \quad x \in D \quad \text{(B.26)} \\
\nu^{\text{exact}}(s^k) &= \sum_{x \in S^k} \nu^{\text{aux}}(x') \quad 1 \leq k \leq K \quad \text{(B.27)} \\
\nu^{\text{exact}}(s^k) &= \sum_{x \in S^k} \nu^{\text{aux}}(x) \quad K + 1 \leq k \leq N.
\end{align*}
$$

\(^{11}\) The auxiliary DTHMC is unichain (cf. Footnote 9) and all states $x \in S$ are recurrent for the auxiliary DTHMC (cf. Footnote 10). On the other hand, some states $x \in D'$ may be transient for the auxiliary DTHMC. However, for each $k$ with $1 \leq k \leq K$, the set $\{x' \in D' \mid x \in S^k\}$ contains at least one recurrent state (cf. Footnote 10). Using that transient states $x \in D'$ satisfy $\nu^{\text{aux}}(x) = 0$, Theorem A.1.5 is applied to the restriction of the auxiliary DTHMC to its recurrent states.
Together with the results of Lemma B.2.7, we obtain the first and third equation of the lemma. For the second equation, consider (for \(1 \leq k \leq K\))

\[
\nu^\text{exact}(s^k) + \sum_{x \in S^k} \nu^\text{exact}(x) = \sum_{x \in S^k} \nu^\text{aux}(x') + \sum_{x \in S^k} \nu^\text{aux}(x)
\]

\[
= \sum_{x \in S^k} \left( \nu^\text{aux}(x') + \nu^\text{aux}(x) \right). \tag{B.27}
\]

\[= \sum_{x \in S^k} (\xi(x)). \quad \text{(Lemma B.2.7)} \]

\[\square \text{(Lemma B.2.8)} \]

The following theorem establishes part (c). Recall that \(\lambda_{\tilde{M}_{\text{rec}}, \mu}\) denotes the average expected cost of \(\tilde{M}_K\) corresponding to \(\mu\).

**Theorem B.2.9** Assume

\[\tilde{g}(s^k) \geq \max_{x \in S^k} \tilde{g}(x, \mu(x)) \quad 1 \leq k \leq N.\]

Let \(\lambda^\text{exact}_{\tilde{M}_K, \mu}\) denote the average expected cost of \(\tilde{M}_K\) corresponding to \(\mu\) with transition probabilities \(\tilde{p}^\text{exact}_\mu(s^l | s^k)\) in aggregated states. Then

\[\lambda^\text{exact}_{\tilde{M}_K, \mu} \geq \lambda_{\tilde{M}_{\text{rec}}, \mu}.\]

We first prove the following lemma. Recall that \(\xi\) denotes the stationary distribution of \(\tilde{X}_\mu\) (the embedded Markov chain of \(\tilde{M}_{\text{rec}}\) corresponding to \(\mu\)) and that \(\nu\) denotes the stationary distribution of \(\tilde{X}_{K, \mu}\) (the embedded Markov chain of \(\tilde{M}_K\) corresponding to \(\mu\)).

**Lemma B.2.10** Assume \(\nu(x^0) \leq \xi(x^0)\). Then

\[\nu(x) \leq \xi(x) \quad x \in D.\]

From Lemma B.2.8, we have \(\nu^\text{exact}(x^0) = \xi(x^0)\) and thus, by Lemma B.2.10, \(\nu^\text{exact}(x) \leq \xi(x)\) for all \(x \in \{x^0\} \cup D\). This is also proven in [LM94, Section 3].
Proof of Lemma B.2.10: Due to definition (B.18), we have to show that \( \nu_D \leq \xi_D \). For \( x \in S \), let \( p_x, \in \mathbb{R}^{1 \times |S|} \) with \( (p_x,)_x := \hat{p}(\bar{x} \mid x, \mu(x)) \) and let \( p_x,|D \) denote the subvector of \( p_x, \) corresponding to the set of states \( D \). From the condition \( \xi^T = \xi^T P_{\mu} \) we obtain (cf. (B.20))

\[
\xi_D = \xi(x^0)p_0^T + P_D^T \xi_D + \sum_{x \in A} \xi(x)(p_x,|D)^T,
\]
i.e.,

\[
(I - P_D^T)\xi_D = \xi(x^0)p_0^T + \sum_{x \in A} \xi(x)(p_x,|D)^T.
\]

By Lemma B.2.6, we have

\[
\xi_D \geq (I - P_D^T)^{-1} \nu(x^0)p_0^T \geq 0.
\]

\[(B.19)\]

\[
\nu_D \geq (I - P_D^T)^{-1} \nu(x^0)p_0^T.
\]

\[\square\text{(Lemma B.2.10)}\]

Proof of Theorem B.2.9: From Lemma B.2.8 we have \( \nu^{\text{exact}}(x^0) = \xi(x^0) \). Thus, by applying Lemma B.2.10 to \( \hat{X}^{\text{exact}}_{\mu} \), we can write \( \xi(x) = \nu^{\text{exact}}(x) + \epsilon(x) \) with \( \epsilon(x) \geq 0 \) for all \( x \in \bigcup_{k=0}^K S^k \) (obviously, \( \epsilon(x^0) = 0 \)). From Lemma B.2.8, we have for \( 1 \leq k \leq K \)

\[
\nu^{\text{exact}}(s^k) + \sum_{x \in S^k} \nu^{\text{exact}}(x) = \sum_{x \in S^k} \xi(x) = \sum_{x \in S^k} \nu^{\text{exact}}(x) + \sum_{x \in S^k} \epsilon(x),
\]

and thus

\[
\nu^{\text{exact}}(s^k) = \sum_{x \in S^k} \epsilon(x).
\]

(B.28)
Now consider
\[
\lambda_{\tilde{M}_{\text{rec}}, \mu} \overset{(5.34)}{=} \sum_{x \in S_{\text{recurrent}}} \xi(x) \tilde{g}(x, \mu(x))
\]
\[
\begin{align*}
&= \sum_{x \in \bigcup_{k=0}^{K} S^k} \xi(x) \tilde{g}(x, \mu(x)) + \sum_{k=K+1}^{N} \sum_{x \in S^k} \xi(x) \bar{g}(x, \mu(x)) \\
&\leq \sum_{x \in \bigcup_{k=0}^{K} S^k} (\nu^{\text{exact}}(x) + \epsilon(x)) \tilde{g}(x, \mu(x)) + \sum_{k=K+1}^{N} \bar{g}(s^k) \sum_{x \in S^k} \xi(x) \\
&= \sum_{x \in \bigcup_{k=0}^{K} S^k} \nu^{\text{exact}}(x) \tilde{g}(x, \mu(x)) + \sum_{k=0}^{K} \sum_{x \in S^k} \epsilon(x) \bar{g}(x, \mu(x)) \\
&\quad + \sum_{k=K+1}^{N} \bar{g}(s^k) \nu^{\text{exact}}(s^k) \\
&\leq \sum_{x \in \bigcup_{k=0}^{K} S^k} \nu^{\text{exact}}(x) \tilde{g}(x, \mu(x)) + \sum_{k=1}^{K} \sum_{x \in S^k} \epsilon(x) \bar{g}(x, \mu(x)) \\
&\quad + \sum_{k=K+1}^{N} \bar{g}(s^k) \nu^{\text{exact}}(s^k) \\
&= \sum_{x \in \bigcup_{k=0}^{K} S^k} \nu^{\text{exact}}(x) \tilde{g}(x, \mu(x)) + \sum_{k=1}^{N} \nu^{\text{exact}}(s^k) \tilde{g}(s^k) \\
&\overset{(5.35)}{=} \lambda^{\text{exact}}_{\tilde{M}, \mu}. 
\end{align*}
\]
\(\square\) (Theorem B.2.9)
“Modification” (part (e) of the proof):

In the following, we consider two (a priori arbitrary) sets of transition probabilities $p_{k,l}$ and $\hat{p}_{k,l}$ in the aggregated states $s^1, \ldots, s^N$ of $\tilde{X}_{K,\mu}$. The next theorem provides conditions on $p_{k,l}$ and $\hat{p}_{k,l}$ and on the cost $\tilde{g}(s^k)$ in aggregated states which assure that the average expected cost (of $\tilde{M}_K$) corresponding to $\hat{p}_{k,l}$ provide an upper bound on the average expected cost corresponding to $p_{k,l}$. To establish part (e), $\hat{p}_{k,l}$ “play the role” of $\tilde{p}(s^l \mid s^k)$ (as specified in Theorem 5.3.4) and $p_{k,l}$ “play the role” of $\tilde{p}_\mu^\text{exact}(s^l \mid s^k)$ (of the “exact aggregation”), cf. Remark B.2.12 below.

**Theorem B.2.11** Assume the transition probabilities $p_{k,l}$ and $\hat{p}_{k,l}$ in the aggregated states $s^1, \ldots, s^N$ satisfy

$$0 < p_{k,k+1} \leq \hat{p}_{k,k+1} \quad 1 \leq k \leq N - 1 \quad (B.29)$$

$$p_{k,k-1} \geq \hat{p}_{k,k-1} > 0 \quad 1 \leq k \leq N \quad (B.30)$$

$$\hat{p}_{k,k+1} \geq \max_{x \in S^k} \sum_{\bar{x} \in S^{k+1}} \tilde{p}(\bar{x} \mid x, \mu(x)) \quad 1 \leq k \leq N - 1 \quad (B.31)$$

$$\hat{p}_{k,k-1} \leq \min_{x \in S^k} \sum_{\bar{x} \in S^{k-1}} \tilde{p}(\bar{x} \mid x, \mu(x)) \quad 1 \leq k \leq N \quad (B.32)$$

Further assume that

$$\tilde{g}(s^k) \geq \max_{x \in S^k} \tilde{g}(x, \mu(x)) \quad 1 \leq k \leq N,$$

and

$$\tilde{g}(s^{k+1}) \geq \tilde{g}(s^k) \geq \tilde{g}(x^0, u^0) \quad 1 \leq k \leq N - 1.$$

Let $\lambda_\mu$ and $\hat{\lambda}_\mu$ denote the average expected cost of $\tilde{M}_K$ corresponding to $\mu$ and with transition probabilities $p_{k,l}$ and $\hat{p}_{k,l}$, respectively, in aggregated states. Then

$$\hat{\lambda}_\mu \geq \lambda_\mu.$$ 

**Remark B.2.12** Assume that the transition probabilities $\tilde{p}(s^l \mid s^k)$ of $\tilde{M}_K$ satisfy the assumptions in Theorem 5.3.4. From (B.13) and (B.14)
and the definition of \( \hat{p}^\text{exact}_\mu(s^l | s^k) \) we observe
\[
\hat{p}^\text{exact}_\mu(s^{k+1} | s^k) \leq \hat{p}(s^{k+1} | s^k) \quad 1 \leq k \leq N - 1
\]
\[
\hat{p}^\text{exact}_\mu(s^{k-1} | s^k) \geq \hat{p}(s^{k-1} | s^k) \quad 1 \leq k \leq N.
\]
Thus, as \( \hat{p}^\text{exact}_\mu(s^{k+1} | s^k) > 0 \) (due to Footnote 10 and Lemma B.2.3 from the proof of Theorem 5.3.3; cf. above) and as \( \hat{p}(s^{k-1} | s^k) > 0 \) (by assumption), \( \hat{p}^\text{exact}_\mu(s^l | s^k), \hat{p}(s^l | s^k) \) satisfy (B.29) and (B.30) (\( \hat{p}(s^l | s^k) \) “play the role” of \( \hat{p}_{k,l} \)). Further, (B.31) and (B.32) are assumptions of Theorem 5.3.4.

Recall that \( \lambda_{\tilde{M}_K,\mu} \) denotes the average expected cost of \( \tilde{M}_K \) corresponding to the deterministic stationary policy \( \mu \) and with transition probabilities \( \hat{p}(s^l | s^k) \). From Theorem B.2.11 we obtain
\[
\lambda_{\tilde{M}_K,\mu} \geq \lambda_{\tilde{M}_K,\mu}^\text{exact},
\]
which establishes part (e) of the proof.

The key property for proving Theorem B.2.11 is given in the following lemma. Let \( \nu \) and \( \hat{\nu} \) denote the stationary distributions of \( \tilde{X}_K,\mu \) with transition probabilities \( p_{k,l} \) and \( \hat{p}_{k,l} \), respectively, in aggregated states. The total stationary probability of each “slice” \( S^k \cup s^k \) is given by (cf. (5.36)):
\[
\nu^+(k) := \begin{cases} 
\nu(x^0) & \text{if } k = 0 \\
 e^T \nu s_k + \nu(s^k) & \text{if } 1 \leq k \leq K \\
 \nu(s^k) & \text{if } K + 1 \leq k \leq N .
\end{cases}
\]
\( \hat{\nu}^+(k), 0 \leq k \leq N, \) is defined analogously with \( \hat{\nu} \) instead of \( \nu \).

**Lemma B.2.13** Assume the transition probabilities \( p_{k,l} \) and \( \hat{p}_{k,l} \) in the aggregated states \( s^1, \ldots, s^N \) satisfy (B.29), (B.30), (B.31), and (B.32). Then, for \( 1 \leq k \leq N - 1 \), we have
\[
\nu^+(k) < \hat{\nu}^+(k) \Rightarrow \nu^+(k + 1) < \hat{\nu}^+(k + 1).
\]

To prove Lemma B.2.13 and Theorem B.2.11, we need the following four lemmas.
Lemma B.2.14 Let $\alpha := \frac{\hat{\nu}(x^0)}{\nu(x^0)}$. Then

$$\hat{\nu}(x) = \alpha \nu(x) \quad x \in D. \quad \text{(B.33)}$$

Proof of Lemma B.2.14: Due to the definition of $\nu_D$ we need to show that $\hat{\nu}_D = \alpha \nu_D$. Consider

$$\hat{\nu}_D^{(B.19)} = \hat{\nu}(x^0)(I - P_D^T)^{-1} p_0^T,$$

$$= \frac{\hat{\nu}(x^0)}{\nu(x^0)} \nu(x^0)(I - P_D^T)^{-1} p_0^T \quad \text{(B.19)}.$$

Lemma B.2.15 Consider a DTHMC $X$ with state space $\{0, 1, \ldots, N\}$ and transitions as in Figure B.4 (transitions are only allowed back to the state itself or to neighboring states, except for state 0, where the outgoing transition goes to state $K + 1$). We consider two sets of transition probabilities $p_{k,l} := p(l \mid k)$ and $\hat{p}_{k,l} := \hat{p}(l \mid k)$ for this Markov chain. Assume these transition probabilities satisfy

$$0 < p_{k,k+1} \leq \hat{p}_{k,k+1} \quad 1 \leq k \leq N - 1,$$

$$p_{k,k-1} \geq \hat{p}_{k,k-1} > 0 \quad 1 \leq k \leq N,$$

$$0 < p_{0,K+1} \leq \hat{p}_{0,K+1}.$$

Let $\pi$, $\hat{\pi}$ denote the (unique$^{12}$) stationary distributions of $X$ corresponding to $p_{k,l}$ and $\hat{p}_{k,l}$, respectively. Then

$$\pi(0) \geq \hat{\pi}(0).$$

A more general version of Lemma B.2.15 is formulated and proven in [MdSeSG89]. For the sake of completeness, we provide the proof of Lemma B.2.15 below. The proof closely follows the proof of [MdSeSG89].

$^{12}$ Due to the positivity of the transition probabilities, $X$ is irreducible and all states are recurrent. Thus, uniqueness of $\pi$, $\hat{\pi}$ follows from Theorem 4.1.4.
Proof of Lemma B.2.15: Applying Theorem A.1.4 with $S'(a)$, $S'(b)$, and $S'(c)$ as indicated in Figure B.5, respectively, yields (note that the last equation is only valid if $K < N - 1$)

\[
\begin{align*}
\pi(1)p_{1,0} &= \pi(0)p_{0,K+1} \\
\pi(k + 1)p_{k+1,k} &= \pi(k)p_{k,k+1} + \pi(0)p_{0,K+1} & 1 \leq k \leq K \\
\pi(k + 1)p_{k+1,k} &= \pi(k)p_{k,k+1} & K + 1 \leq k \leq N - 1.
\end{align*}
\]

Thus

\[
\begin{align*}
\pi(1) &= \frac{p_{0,K+1}}{p_{1,0}} \pi(0) \\
\pi(k + 1) &= \frac{p_{k+1,k}}{p_{k,k+1}} \pi(k) + \frac{p_{0,K+1}}{p_{k+1,k}} \pi(0) & 1 \leq k \leq K \\
\pi(k + 1) &= \frac{p_{k+1,k}}{p_{k,k+1}} \pi(k) & K + 1 \leq k \leq N - 1.
\end{align*}
\]
Clearly, the same formulas hold for $\hat{p}_{k,l}$ and $\hat{\pi}$. Due to the assumptions of the lemma we have

\begin{align*}
\pi(1) &\leq \frac{\hat{p}_{0,K+1}}{\hat{p}_{1,0}} \pi(0) \quad \text{(B.34)} \\
\pi(k + 1) &\leq \frac{\hat{p}_{k,k+1}}{\hat{p}_{k+1,k}} \pi(k) + \frac{\hat{p}_{0,K+1}}{\hat{p}_{k+1,k}} \pi(0) \quad 1 \leq k \leq K \quad \text{(B.35)} \\
\pi(k + 1) &\leq \frac{\hat{p}_{k,k+1}}{\hat{p}_{k+1,k}} \pi(k) \quad K + 1 \leq k \leq N - 1. \quad \text{(B.36)}
\end{align*}

Assume $\pi(0) < \hat{\pi}(0)$. Then, from (B.34), we conclude $\pi(1) < \hat{\pi}(1)$. Applying (B.35) repeatedly for $1 \leq k \leq K$ yields $\pi(k + 1) < \hat{\pi}(k + 1)$ for $1 \leq k \leq K$. Finally (if $K < N - 1$), applying (B.36) repeatedly for $K + 1 \leq k \leq N - 1$ yields $\pi(k + 1) < \hat{\pi}(k + 1)$ for $K + 1 \leq k \leq N - 1$. Hence it was shown that $\pi(k) < \hat{\pi}(k)$ for all $0 \leq k \leq N$. But then

$$1 = \sum_{k=0}^{N} \pi(k) < \sum_{k=0}^{N} \hat{\pi}(k) = 1,$$

which is a contradiction. Thus $\pi(0) \geq \hat{\pi}(0)$. \hfill \Box (Lemma B.2.15)

**Lemma B.2.16** Assume the transition probabilities $p_{k,l}$ and $\hat{p}_{k,l}$ in the aggregated states $s^1, \ldots, s^N$ satisfy (B.29) and (B.30). Then

$$\nu(x^0) \geq \hat{\nu}(x^0).$$

Combining Lemma B.2.16 with Lemma B.2.14 yields $\hat{\nu}(x) \leq \nu(x)$ for all $x \in \{x^0\} \cup D$. This result is also stated in [LM94, Theorem 3.1] and proven in [LM94, Lemma A.2].

**Proof of Lemma B.2.16:** Aggregate all states in $\bigcup_{k=0}^{K} S^k$ into the aggregate state $G$, cf. Figure B.4 on page 219 ($G$ corresponds to the state 0). The transition probabilities from and to the aggregate state $G$ are defined by (cf. Theorem A.1.5)$^{13}$

$$p_{G,K+1} := \frac{(\nu_{S^k})^T p_{s^k,s^{k+1}}}{e^T \nu_{S^k}}$$

$$p_{1,G} := p_{1,0}$$

$^{13}$ Here, we explicitly use (5.24) from page 137.
\[ \hat{p}_{G,K+1} := \frac{(\hat{\nu}_S^k)^T p_{\cdot,s^k+1}}{e^T \hat{\nu}_S^k} \]
\[ \hat{p}_{1,G} := \hat{p}_{1,0}. \]

Let \( \nu^{agg}, \hat{\nu}^{agg} \) denote the corresponding stationary distributions of the (exact) aggregation. Then we have from Theorem A.1.5

\[ \nu^{agg}(G) = \sum_{x \in \bigcup_{k=0}^K S^k} \nu(x) \]
\[ \hat{\nu}^{agg}(G) = \sum_{x \in \bigcup_{k=0}^K S^k} \hat{\nu}(x). \]

We also have

\[ \hat{\nu}^{agg}(G) = \sum_{x \in \bigcup_{k=0}^K S^k} \hat{\nu}(x) \quad \text{(Lemma B.2.14)} \]
\[ = \alpha \cdot \sum_{x \in \bigcup_{k=0}^K S^k} \nu(x) = \alpha \cdot \nu^{agg}(G), \]

and thus

\[ \alpha = \frac{\hat{\nu}^{agg}(G)}{\nu^{agg}(G)}. \]

We further observe

\[ \hat{p}_{G,K+1} = \frac{(\hat{\nu}_S^k)^T p_{\cdot,s^k+1}}{e^T \hat{\nu}_S^k} \quad \text{(Lemma B.2.14)} = \frac{\alpha (\nu_S^k)^T p_{\cdot,s^k+1}}{\alpha e^T \nu_S^k} \]
\[ = \frac{(\hat{\nu}_S^k)^T p_{\cdot,s^k+1}}{e^T \nu_S^k} = p_{G,K+1}, \]

and \( p_{1,G} \geq \hat{p}_{1,G} \) (by assumption (B.30)).

Hence, from Lemma B.2.15 we have \( \hat{\nu}^{agg}(G) \leq \nu^{agg}(G) \), i.e., \( \alpha \leq 1 \) and thus (by definition of \( \alpha \) in Lemma B.2.14) \( \hat{\nu}(x^0) \leq \nu(x^0) \).

\( \square \) (Lemma B.2.16)
In the following we write $\nu_k$ for $\nu(s^k)$ for $1 \leq k \leq N$.

**Lemma B.2.17** Assume $p_{k+1,k} > 0$ for $1 \leq k \leq N - 1$. Then

\[
\nu_{k+1} = \frac{p_{k,k+1}}{p_{k+1,k}} \nu_k \quad K + 1 \leq k \leq N - 1 \quad (B.37)
\]

\[
\nu_{k+1} = \frac{p_{k,k+1}}{p_{k+1,k}} \nu_k + \frac{1}{p_{k+1,k}} (\nu_{s^k})^T p_{s^{k+1},s^k} \quad 1 \leq k \leq K \quad (B.38)
\]

Clearly, if $\hat{p}_{k+1,k} > 0$ for $1 \leq k \leq N - 1$, the same formulas hold for $\hat{\nu}_k$ with $\hat{p}_{k,l}$ instead of $p_{k,l}$.

**Proof of Lemma B.2.17:**

Note that for $K = N - 1$, (B.37) is “invalid”, i.e., there is no equation of type (B.37). Applying Theorem A.1.4 with $S'_1$ (for $K + 1 \leq k \leq N - 1$ if such $k$ exist) and $S'_2$ (for $1 \leq k \leq K$) as indicated in Figure B.6 yields\(^\dagger\)

\[
\nu_{k+1} p_{k+1,k} = p_{k,k+1} \nu_k \quad K + 1 \leq k \leq N - 1
\]

\[
\nu_{k+1} p_{k+1,k} = p_{k,k+1} \nu_k + (\nu_{s^k})^T p_{s^{k+1},s^k} \quad 1 \leq k \leq K
\]

**Figure B.6** Total “flow” in equals total “flow” out.

\(\square\) (Lemma B.2.17)

**Proof of Lemma B.2.13:** Let $\alpha := \frac{\hat{\nu}(x^0)}{\hat{\nu}(x)}$. From Lemma B.2.14 we have

\[
\hat{\nu}(x) = \alpha \nu(x) \quad x \in \bigcup_{k=1}^{K} S^k, \quad (B.39)
\]

and from Lemma B.2.16 we have

\[
\alpha \leq 1. \quad (B.40)
\]

\(^\dagger\) Here, we explicitly use (5.24) from page 137.
Let $K + 1 \leq k \leq N - 1$ (if such a $k$ exists). Assume $\nu^+(k) < \hat{\nu}^+(k)$, i.e., $\nu_k < \hat{\nu}_k$ (by definition). Using (B.37) of Lemma B.2.17 we obtain

$$\nu_{k+1} \overset{(B.37)}{=} \frac{p_{k,k+1}}{p_{k+1,k}} \nu_k < \frac{\hat{p}_{k,k+1}}{\hat{p}_{k+1,k}} \hat{\nu}_k \overset{(B.37)}{=} \hat{\nu}_{k+1},$$

i.e., $\nu^+(k+1) < \hat{\nu}^+(k+1)$, and the claim is true for $K + 1 \leq k \leq N - 1$.

Let $1 \leq k \leq K$. Assume $\nu^+(k) < \hat{\nu}^+(k)$, i.e., $e^T \nu_{S^k} + \nu_k < e^T \hat{\nu}_{S^k} + \hat{\nu}_k$ (by definition). As $\nu_{S^k} \overset{(B.39)}{=} \frac{1}{\alpha} \hat{\nu}_{S^k}$, this assumption can be reformulated as

$$\nu_k < e^T \hat{\nu}_{S^k} \left(1 - \frac{1}{\alpha}\right) + \hat{\nu}_k.$$ 

Using (B.38) of Lemma B.2.17 we obtain

$$\nu_{k+1} \overset{(B.38)}{=} \frac{p_{k,k+1}}{p_{k+1,k}} \nu_k + \frac{1}{p_{k+1,k}} \left(\nu_{S^k}\right)^T \overline{p}_{.,S^k+1} \leq \frac{\hat{p}_{k,k+1}}{\hat{p}_{k+1,k}} \left(e^T \hat{\nu}_{S^k} \left(1 - \frac{1}{\alpha}\right) + \hat{\nu}_k\right) + \frac{1}{\hat{p}_{k+1,k}} \left(\nu_{S^k}\right)^T \overline{p}_{.,S^k+1}$$

$$= \frac{\hat{p}_{k,k+1}}{\hat{p}_{k+1,k}} \left(1 - \frac{1}{\alpha}\right) e^T \hat{\nu}_{S^k} + \frac{\hat{p}_{k,k+1}}{\hat{p}_{k+1,k}} \hat{\nu}_k + \frac{1}{\hat{p}_{k+1,k}} \left(\nu_{S^k}\right)^T \overline{p}_{.,S^k+1}$$

$$= \frac{\hat{p}_{k,k+1}}{\hat{p}_{k+1,k}} \left(1 - \frac{1}{\alpha}\right) e^T \hat{\nu}_{S^k} + \hat{\nu}_{k+1} + \frac{1}{\hat{p}_{k+1,k}} \left(\nu_{S^k}\right)^T \overline{p}_{.,S^k+1} \left(-1 + \frac{1}{\alpha}\right)$$

$$= \hat{\nu}_{k+1} + \left(-\frac{\hat{p}_{k,k+1}}{\hat{p}_{k+1,k}} e^T \hat{\nu}_{S^k} + \left(\nu_{S^k}\right)^T \overline{p}_{.,S^k+1}\right).$$

First we consider the special case $k = K$. In this case (B.41) reads

$$\nu_{K+1} < \hat{\nu}_{K+1} + \left(-\frac{\hat{p}_{K,K+1}}{\hat{p}_{K+1,K}} e^T \hat{\nu}_{S^K} + \left(\nu_{S^K}\right)^T \overline{p}_{.,S^K+1}\right)$$

$$= \hat{\nu}_{K+1} + \left(-1\right) \frac{1}{\hat{p}_{K+1,K}} \left(\nu_{S^K}\right)^T \left(p_{.,S^K+1} - \hat{p}_{K,K+1} \cdot e\right).$$
For \( x \in S^K \), we have (5.26)
\[
(5.26) \quad (p_{s_k^{K+1}})_x = \sum_{\bar{x} \in S^{K+1}} \tilde{p}(\bar{x} \mid x, \mu(x)) \leq \hat{p}_{K,K+1},
\]
i.e., \( p_{s_k^{K+1}} - \hat{p}_{K,K+1} \cdot e \leq 0 \). As \( \frac{1}{\alpha} \geq 1 \), we obtain
\[
\nu_{K+1} < \hat{\nu}_{K+1},
\]
i.e., \( \nu^+(K + 1) < \hat{\nu}^+(K + 1) \), and the claim is true for \( k = K \).

Now let \( 1 \leq k \leq K - 1 \). As \( \nu_{S_k^{K+1}} \overset{(B.39)}{=} \frac{1}{\alpha} \hat{\nu}_{S_k^{K+1}} \) we obtain with (B.41)
\[
e^T \nu_{S_k^{K+1}} + \nu_{k+1} = \frac{1}{\alpha} e^T \hat{\nu}_{S_k^{K+1}} + \nu_{k+1}
\overset{(B.41)}{<} \frac{1}{\alpha} e^T \hat{\nu}_{S_k^{K+1}} + \hat{\nu}_{k+1}
\]
\[
+ \left( \frac{1}{\alpha} - 1 \right) \frac{1}{\hat{p}_{k+1,k}} \left( -\hat{p}_{k,k+1} e^T \hat{\nu}_{S_k} + (\hat{\nu}_{S_K})^T p_{s_k^{K+1}} \right)
\]
\[
= \ldots + e^T \hat{\nu}_{S_k^{K+1}} - e^T \hat{\nu}_{S_k^{K+1}} = 0
\]
\[
= e^T \hat{\nu}_{S_k^{K+1}} + \hat{\nu}_{k+1} + \left( \frac{1}{\alpha} - 1 \right) \frac{1}{\hat{p}_{k+1,k}}
\]
\[
\cdot \left( \hat{p}_{k+1,k} e^T \hat{\nu}_{S_k^{K+1}} - \hat{p}_{k,k+1} e^T \hat{\nu}_{S_k} + (\hat{\nu}_{S_K})^T p_{s_k^{K+1}} \right)
\overset{=}=: \gamma_k
\]

Using Theorem A.1.4, we obtain for \( 1 \leq k \leq K - 1 \) (cf. Figure B.7):\(^{15}\)

\[\text{Figure B.7} \quad \text{Total "flow" in equals total "flow" out.}\]

\(^{15}\) Here, we explicitly use (5.24) from page 137.
i.e.,
\[ \hat{p}_{k+1,k} e^T \hat{\nu}_{S^{k+1}} + (\hat{\nu}_{S^K})^T \hat{p}_{-,s^{K+1}} \leq \hat{p}_{k,k+1} e^T \hat{\nu}_{S^{k}}, \]
and thus \( \hat{\gamma}_k \leq 0 \).

As \( \frac{1}{\alpha} \geq 1 \) we conclude
\[ e^T \nu_{S^{k+1}} + \nu_{k+1} < e^T \hat{\nu}_{S^{k+1}} + \hat{\nu}_{k+1}, \]
i.e., \( \nu^+(k+1) < \hat{\nu}^+(k+1) \), and the claim also holds for \( 1 \leq k \leq K - 1 \).

\( \square \) (Lemma B.2.13)

**Proof of Theorem B.2.11:** From Lemma B.2.16 we have
\[ \nu(x^0) \geq \hat{\nu}(x^0). \quad (B.42) \]

Thus \( \alpha := \frac{\hat{\nu}(x^0)}{\nu(x^0)} \leq 1 \) and we obtain from Lemma B.2.14
\[ \nu(x) \geq \hat{\nu}(x) \quad x \in \bigcup_{k=1}^{K} S^k. \quad (B.43) \]

For notational convenience, define
\[ \tilde{g}(s^0) := \tilde{g}(x^0, u^0), \]
i.e., \( \tilde{g}(s^0) = \tilde{g}(x^0, \hat{\mu}(x^0)) \) for any deterministic stationary policy \( \hat{\mu} \in \Pi^{MD} \) (cf. (5.30)). Consider

\[
\lambda_\mu - \hat{\lambda}_\mu \overset{(5.35)}{=} \sum_{0 \leq k \leq K} \sum_{x \in S^k} (\nu(x) - \hat{\nu}(x)) \tilde{g}(x, \mu(x)) + \sum_{1 \leq k \leq N} (\nu_k - \hat{\nu}_k) \tilde{g}(s^k)
\]

\[
\leq (\nu(x^0) - \hat{\nu}(x^0)) \tilde{g}(x^0, \mu(x^0))
\]

\[
+ \sum_{1 \leq k \leq K} \tilde{g}(s^k) (\sum_{x \in S^k} \nu(x) + \nu_k - (\sum_{x \in S^k} \hat{\nu}(x) + \hat{\nu}_k))
\]

\[
+ \sum_{K < k \leq N} \tilde{g}(s^k) (\nu_k - \hat{\nu}_k)
\]

\[
= \sum_{0 \leq k \leq N} (\nu^+(k) - \hat{\nu}^+(k)) \tilde{g}(s^k). \tag{B.44}
\]

From (B.42) we have \( \nu^+(0) - \hat{\nu}^+(0) \geq 0 \).

**Case 1:** \( \nu^+(k) - \hat{\nu}^+(k) \geq 0 \) for all \( 0 \leq k \leq N \). Assume there exists \( \bar{k} \), \( 0 \leq \bar{k} \leq N \), such that \( \nu^+(\bar{k}) - \hat{\nu}^+(\bar{k}) > 0 \). Then we have

\[
1 = \sum_{k=0}^{N} \nu^+(k) + \nu^+(\bar{k}) > \sum_{k=0}^{N} \hat{\nu}^+(k) = 1,
\]

which is a contradiction. Hence \( \nu^+(k) = \hat{\nu}^+(k) \) for all \( 0 \leq k \leq N \). From (B.44) we obtain

\[
\lambda_\mu - \hat{\lambda}_\mu \leq 0.
\]

**Case 2:** There exists \( 1 \leq k \leq N \) such that \( \nu^+(k) - \hat{\nu}^+(k) < 0 \). Define

\[
k_0 := \min\{k \mid \nu^+(k) - \hat{\nu}^+(k) < 0\}.
\]
Thus, $\nu^+(k) - \hat{\nu}^+(k) \geq 0$ for all $0 \leq k < k_0$ and (due to Lemma B.2.13) $\nu^+(k) - \hat{\nu}^+(k) < 0$ for all $k_0 \leq k \leq N$. From (B.44) we obtain

$$\lambda_{\mu} - \hat{\lambda}_{\mu} \leq \sum_{0 \leq k \leq N} (\nu^+(k) - \hat{\nu}^+(k)) \tilde{g}(s^k)$$

$$= \sum_{0 \leq k < k_0} (\nu^+(k) - \hat{\nu}^+(k)) \tilde{g}(s^k) + \sum_{k_0 \leq k \leq N} (\nu^+(k) - \hat{\nu}^+(k)) \tilde{g}(s^k)$$

$$\leq \tilde{g}(s^{k_0}) \sum_{0 \leq k \leq N} (\nu^+(k) - \hat{\nu}^+(k))$$

$$= \tilde{g}(s^{k_0})(\sum_{0 \leq k \leq N} \nu^+(k) - \sum_{0 \leq k \leq N} \hat{\nu}^+(k)) = 0.$$}

□(Theorem B.2.11)

---

**Theorem 5.3.6 (on page 143)** Let $n_{\text{res}} \geq 1$. Let $\sigma$, $\hat{\sigma}$ denote permutations of the set $\{1, \ldots, |C|\}$ that satisfy

$$\alpha^\text{fail}_{c_{\sigma(1)}} \leq \ldots \leq \alpha^\text{fail}_{c_{\sigma(|C|)}},$$

and

$$\alpha^\text{repair}_{c_{\hat{\sigma}(1)}} \leq \ldots \leq \alpha^\text{repair}_{c_{\hat{\sigma}(|C|)}},$$

respectively. For $0 \leq k \leq N^{\max} - 1$, we define (using the convention that $\sum_{i=a}^{b} x = 0$ whenever $b < a$)\(^{16}\)

$$\bar{l}(k) := \max\{\bar{l} \mid \sum_{l=\bar{l}}^{|C|} m_{c_{\sigma(l)}} \geq N^{\max} - k\}$$

$$\bar{m}(k) := N^{\max} - k - \sum_{l=\bar{l}(k)+1}^{|C|} m_{c_{\sigma(l)}},$$

---

\(^{16}\) The idea of $\bar{l}(k)$ is that $\{c_{\sigma(\bar{l}(k))}, \ldots, c_{\sigma(|C|)}\}$ is the smallest set of components with the largest failure rates that maximally require at least $N^{\max} - k$ resources. The idea of $\bar{m}(k)$ is that $\{c_{\hat{\sigma}(1)}, \ldots, c_{\hat{\sigma}(\bar{l}(k))}\}$ is the smallest set of components with the smallest repair rates that maximally require at least $\min\{k, n_{\text{res}}\}$ resources. $\bar{l}(k)$, $\bar{m}(k)$, $\bar{l}(k)$, and $\bar{m}(k)$ are illustrated in Figure 5.9.
and for $1 \leq k \leq N_{\text{max}}$, we define

\[
\tilde{l}(k) := \min \{ \tilde{l} \mid \sum_{l=1}^{\tilde{l}} m_{c_{\sigma(l)}} \geq \min \{ k, n_{\text{res}} \} \}
\]

\[
\overline{m}(k) := \min \{ k, n_{\text{res}} \} - \sum_{l=1}^{\tilde{l}(k)-1} m_{c_{\sigma(l)}}.
\]

Define the transition probabilities $\tilde{p}(s^l \mid s^k)$ in aggregated states as follows:

\[
\tilde{p}(s^{k+1} \mid s^k) := \frac{1}{\eta_{\mathcal{M}}} \cdot q_{k}^{\text{UB,+}} \quad 1 \leq k \leq N - 1
\]

\[
\tilde{p}(s^{k-1} \mid s^k) := \frac{1}{\eta_{\mathcal{M}}} \cdot q_{k}^{\text{UB,-}} \quad 1 \leq k \leq N
\]

\[
\tilde{p}(s^l \mid s^k) := 0 \quad 1 \leq k \leq N, |l - k| > 1
\]

\[
\tilde{p}(s^k \mid s^k) := 1 - \tilde{p}(s^{k-1} \mid s^k) - \tilde{p}(s^{k+1} \mid s^k) \quad 1 \leq k \leq N - 1
\]

\[
\tilde{p}(s^N \mid s^N) := 1 - \tilde{p}(s^{N-1} \mid s^N),
\]

where

\[
q_{k}^{\text{UB,+}} := \sum_{l=\tilde{l}(k)+1}^{\lvert \mathcal{C} \rvert} m_{c_{\sigma(l)}} \cdot \alpha_{c_{\sigma(l)}}^{\text{fail}} + \overline{m}(k) \cdot \alpha_{c_{\sigma(\tilde{l}(k))}}^{\text{fail}} \quad 1 \leq k \leq N_{\text{max}} - 1,
\]

\[
q_{k}^{\text{UB,-}} := \sum_{l=1}^{\tilde{l}(k)-1} m_{c_{\sigma(l)}} \cdot \alpha_{c_{\sigma(l)}}^{\text{repair}} + \overline{m}(k) \cdot \alpha_{c_{\sigma(\tilde{l}(k))}}^{\text{repair}} \quad 1 \leq k \leq N_{\text{max}}.
\]

Then, $\tilde{p}(\cdot \mid s^k), 1 \leq k \leq N$, are indeed probability distributions and the transition probabilities $\tilde{p}(s^l \mid s^k)$ satisfy the assumptions of Theorem 5.3.4.

**Proof of Theorem 5.3.6:**

The proof that $\tilde{p}(\cdot \mid s^k), 1 \leq k \leq N$, are probability distributions is given at the end. Clearly, (5.27) is satisfied. From the definitions of $\overline{m}(k)$ and $\tilde{l}(k)$ we have for $0 \leq k \leq N_{\text{max}} - 1$

\[
\sum_{l=\tilde{l}(k)+1}^{\lvert \mathcal{C} \rvert} m_{c_{\sigma(l)}} + \overline{m}(k) = N_{\text{max}} - k \geq 1. \quad (B.45)
\]
Hence, $q_k^{UB,+}$ sums at least one failure rate and as all failure rates are 
positive, (5.28) is satisfied. From the definitions of $\overline{m}(k)$ and $\underline{l}(k)$ we 
have for $1 \leq k \leq N^{\text{max}}$
\[
\sum_{l=1}^{\underline{l}(k)-1} m_{c_{\sigma(l)}} + m(k) = \min\{k, n_{\text{res}}\} \geq 1. \quad (B.46)
\]
Hence, $q_k^{UB,-}$ sums at least one repair rate (which are all positive) and 
(5.29) holds.

We consider the continuous-time MDP $\mathcal{M}$, restricted to the set of states 
$S^{\text{recurrent}}$. The (restricted) jump probabilities $p(\cdot \mid x, u)$ are still probabil-
ity distributions (because $S^{\text{recurrent}}$ is closed for the embedded Markov 
chain $\tilde{X}_\mu$ of $\tilde{M}$ corresponding to any deterministic stationary policy 
$\mu \in \Pi^{\text{MD}}$ (cf. Theorem 5.2.6) and due to (4.19)).

Let $x \in S^k$ ($1 \leq k \leq N$) and $u \in U(x)$. We discuss which transitions 
lead to which set of states $S^l$. As in Section 5.2.3, $T(x, u)$ denotes the set of 
active timers for $x$ and $u$. This set can be partitioned into (cf. (5.24)) $T_{+1}(x, u)$ and $T_{-1}(x, u)$ which denote the sets of timers that, when they 
elapse, lead to transitions to states in $S^{k+1}$ and $S^{k-1}$, respectively. Note 
that for $k = N$, $T_{+1}(x, u) = \emptyset$.

$T_{+1}(x, u)$ contains at most $\sum_{c \in C}(m_c - x_c) = N^{\text{max}} - \sum_{c \in C} x_c$ timers of 
(independent or common mode) failures (cf. Section 5.2.3). As $x \in S^k$, 
we have $\sum_{c \in C} x_c = k$ and thus $|T_{+1}(x, u)| \leq N^{\text{max}} - k$. As $q_k^{UB,+}$ sums 
the largest $N^{\text{max}} - k$ failure rates (cf. (B.45)), we have (for $1 \leq k \leq N - 1$)
\[
\sum_{T \in T_{+1}(x, u)} \alpha_T \leq q_k^{UB,+} \quad x \in S^k, u \in U(x). \quad (B.47)
\]

$T_{-1}(x, u)$ contains $\sum_{c \in C} u_c$ repair timers (cf. Section 5.2.3). Thus 
$|T_{-1}(x, u)| = \sum_{c \in C} u_c = \min\{k, n_{\text{res}}\}$ (using $\sum_{c \in C} x_c = k$ for $x \in S^k$). 
As $q_k^{UB,-}$ sums the smallest $\min\{k, n_{\text{res}}\}$ repair rates, we conclude (for 
$1 \leq k \leq N$)
\[
\sum_{T \in T_{-1}(x, u)} \alpha_T \geq q_k^{UB,-} \quad x \in S^k, u \in U(x). \quad (B.48)
\]
We now consider the uniformization $\tilde{M}_{\text{rec}}$. Let $1 \leq k \leq N - 1$ and let $x \in S^k$ and $u \in U(x)$. Then we have

$$\sum_{\bar{x} \in S^{k+1}} \tilde{p}(\bar{x} | x, u) \overset{(5.11)}{=} \sum_{\bar{x} \in S^{k+1}} \frac{1}{\eta_M} \alpha(\bar{x} | x, u)$$

$$= \frac{1}{\eta_M} \sum_{\bar{x} \in S^{k+1}} \alpha(\bar{x} | x, u)$$

$$\overset{(5.7)}{=} \frac{1}{\eta_M} \sum_{\bar{x} \in S^{k+1} \cap S_{\text{subsequent}}(x,u)} \sum_{T \in T(\bar{x}|x,u)} \alpha_T$$

$$= \frac{1}{\eta_M} \sum_{T \in T_{-1}(x,u)} \alpha_T$$

\[(B.47)\]

$$\leq \frac{1}{\eta_M} q_k^{UB,+} = \tilde{p}(s^{k+1} | s^k),$$

and thus

$$\max_{x \in S^k, u \in U(x)} \sum_{\bar{x} \in S^{k+1}} \tilde{p}(\bar{x} | x, u) \leq \tilde{p}(s^{k+1} | s^k) \quad 1 \leq k \leq N - 1.$$ 

Let $1 \leq k \leq N$ and let $x \in S^k$ and $u \in U(x)$. Analogously to above, we obtain

$$\sum_{\bar{x} \in S^{k-1}} \tilde{p}(\bar{x} | x, u) \overset{(5.11)}{=} \sum_{\bar{x} \in S^{k-1}} \frac{1}{\eta_M} \alpha(\bar{x} | x, u)$$

$$= \frac{1}{\eta_M} \sum_{T \in T_{-1}(x,u)} \alpha_T$$

\[(B.48)\]

$$\geq \frac{1}{\eta_M} q_k^{UB,-} = \tilde{p}(s^{k-1} | s^k),$$

and thus

$$\min_{x \in S^k, u \in U(x)} \sum_{\bar{x} \in S^{k-1}} \tilde{p}(\bar{x} | x, u) \geq \tilde{p}(s^{k-1} | s^k) \quad 1 \leq k \leq N.$$ 

We finally prove that $\tilde{p}(\cdot | s^k), 1 \leq k \leq N$, are probability distributions. Above, we have already argued that (5.28) and (5.29) are satisfied and by definition $\tilde{p}(s^k | s^k) \leq 1$ for $1 \leq k \leq N$. Observe that $m_{c_\sigma(l(k))} \geq \underline{m}(k)$ (for $0 \leq k \leq N^{\max} - 1$) and $m_{c_\sigma(l(k))} \geq \overline{m}(k)$ (for $1 \leq k \leq N^{\max}$) and thus
\[ q_k^{UB,+} < \eta_M \ (0 \leq k \leq \Nmax - 1) \quad \text{and} \quad q_k^{UB,-} < \eta_M \ (1 \leq k \leq \Nmax). \]

Consequently, \( \tilde{p}(s^{k+1} \mid s^k) \leq 1 \) for \( 1 \leq k \leq N - 1(\leq \Nmax - 1) \) and \( \tilde{p}(s^{k-1} \mid s^k) \leq 1 \) for \( 1 \leq k \leq N(\leq \Nmax). \) Also, \( \tilde{p}(s^N \mid s^N) > 0. \) It remains to show that \( \tilde{p}(s^k \mid s^k) \geq 0 \) for \( 1 \leq k \leq N - 1. \) We need the following lemma.

**Lemma B.2.18** Let \( v, w \in \mathbb{R}^n. \) Assume (w.l.o.g.) that \( v_1 \leq v_2 \leq \cdots \leq v_n. \) Let \( \sigma \) be a permutation of the set \( \{1, 2, \ldots, n\} \) such that \( w_{\sigma(1)} \leq w_{\sigma(2)} \leq \cdots \leq w_{\sigma(n)}. \) Let \( \alpha, \beta \in \mathbb{N} \) with \( \alpha + \beta \leq n. \) Then

\[
\Sigma_{\alpha,\beta}(v, w) := \sum_{i=1}^\alpha v_i + \sum_{j=n-\beta+1}^n w_{\sigma(j)} \leq \sum_{i=1}^n \max\{v_i, w_i\},
\]

i.e., the sum of the \( \alpha \) smallest elements of \( v \) and the \( \beta \) largest elements of \( w \) is bounded by the sum of the componentwise maxima of \( v \) and \( w. \)

Before we prove Lemma B.2.18, we discuss its implications. From Lemma B.2.18, we conclude for \( 1 \leq k \leq N - 1 \) \( (\alpha_c^{\text{repair}} \text{ play the role of } v_i, \ \alpha_c^{\text{fail}} \text{ play the role of } w_i, \ \alpha = \min\{k, n_{\text{res}}\} \) and \( \beta = \Nmax - k, \) cf. (B.45) and (B.46))

\[
q_k^{UB,-} + q_k^{UB,+} \leq \sum_{c \in C} m_c \cdot \max\{\alpha_c^{\text{fail}}, \alpha_c^{\text{repair}}\} < \eta_M,
\]

i.e., \( \tilde{p}(s^{k+1} \mid s^k) + \tilde{p}(s^{k-1} \mid s^k) < 1, \) which finishes the proof.

**Proof of Lemma B.2.18:**

Define \( I_\alpha := \{1, 2, \ldots, \alpha\} \) and let \( I_\beta \subseteq \{1, 2, \ldots, n\} \) such that \( \sum_{i \in I_\beta} w_i = \sum_{j=n-\beta+1}^n w_{\sigma(j)} \), i.e.,

\[
\Sigma_{\alpha,\beta}(v, w) = \sum_{i \in I_\alpha} v_i + \sum_{i \in I_\beta} w_i.
\]

Further define \( I_\cap := I_\alpha \cap I_\beta. \) Then, as the sets \( I_\alpha \setminus I_\cap \) and \( I_\beta \setminus I_\cap \) are disjoint, we have

\[
\Sigma_{\alpha,\beta}(v, w) = \sum_{i \in I_\cap} v_i + \sum_{i \in I_\cap} w_i + \sum_{i \in I_\alpha \setminus I_\cap} v_i + \sum_{i \in I_\beta \setminus I_\cap} w_i \\
\leq \sum_{i \in I_\cap} v_i + \sum_{i \in I_\cap} w_i + \sum_{i \in (I_\alpha \setminus I_\cap) \cup (I_\beta \setminus I_\cap)} \max\{v_i, w_i\}.
\]
Assume $I_{\cap} = \emptyset$, i.e., no index $i \in \{1, 2, \ldots, n\}$ appears twice in (B.49). Then,

$$
\Sigma_{\alpha, \beta}(v, w) \leq \sum_{i \in I_{\alpha} \cup I_{\beta}} \max\{v_i, w_i\} \leq \sum_{i=1}^{n} \max\{v_i, w_i\},
$$

and the proof is finished.

In the sequel, we assume that $I_{\cap} \neq \emptyset$. In this case, at least one index appears twice in (B.49). The idea of the proof is to replace the smaller value of $v_i$ and $w_i$ for each index $i$ that appears twice by $\max\{v_{i'}, w_{i'}\}$ of an index $i'$ that does not appear in (B.49).

Let $J := \{1, 2, \ldots, n\} \setminus (I_{\alpha} \cup I_{\beta})$, i.e., $J$ collects those indices which do not appear in (B.49). Observe

$$
|J| = n - |I_{\alpha} \cup I_{\beta}| = n - (|I_{\alpha}| + |I_{\beta}| - |I_{\cap}|) \geq |I_{\cap}|, \quad (B.51)
$$

and

$$
v_i \leq v_j \quad \forall i \in I_{\alpha}, j \in J. \quad (B.52)
$$

For each $i \in I_{\cap}$, choose a distinct $k(i) \in J$ (which is possible due to (B.51)) and define $J_i := \{k(i) \in J \mid i \in I_{\cap}\}$. Further define $I_{\cap, <} := \{i \in I_{\cap} \mid w_i < v_i\}$ and $I_{\cap, \geq} := I_{\cap} \setminus I_{\cap, <}$. As

$$
w_i < v_i \leq v_{k(i)} \quad i \in I_{\cap, <}, \quad (B.53)
$$

we obtain

$$
\begin{align*}
\sum_{i \in I_{\cap}} v_i + \sum_{i \in I_{\cap}} w_i &= \sum_{i \in I_{\cap, <}} v_i + \sum_{i \in I_{\cap, \geq}} v_i \leq \sum_{i \in I_{\cap, <}} v_i + \sum_{i \in I_{\cap, \geq}} v_{k(i)} + \sum_{i \in I_{\cap, \geq}} w_i \\
&\leq \sum_{i \in I_{\cap, <}} v_i + \sum_{i \in I_{\cap}} v_{k(i)} + \sum_{i \in I_{\cap, \geq}} w_i \\
&= \sum_{j \in J} v_j + \sum_{j \in J} w_j \\
&\leq \sum_{i \in I_{\cap}} \max\{v_i, w_i\} + \sum_{j \in J_{\cap}} \max\{v_j, w_j\}. \quad (B.54)
\end{align*}
$$
Hence we conclude (using $\left( I_\alpha \setminus I_{\cap} \right) \cup \left( I_\beta \setminus I_{\cap} \right) = \left( I_\alpha \cup I_\beta \right) \setminus I_{\cap}$)

$\Sigma_{\alpha,\beta}(v, w)^{(B.50),(B.54)} \leq \sum_{i \in I_{\cap}} \max\{v_i, w_i\} + \sum_{j \in J_{\cap}} \max\{v_j, w_j\}
\leq \sum_{j \in J_{\cap}} \max\{v_j, w_j\} + \sum_{i \in (I_\alpha \cup I_\beta) \setminus I_{\cap}} \max\{v_i, w_i\}
\leq \sum_{i \in I_\alpha \cup I_\beta} \max\{v_i, w_i\} + \sum_{j \in J} \max\{v_j, w_j\}
= \sum_{i=1}^{n} \max\{v_i, w_i\}.

□(Lemma B.2.18) □
Bibliography

<table>
<thead>
<tr>
<th>Reference</th>
<th>Title</th>
<th>Publisher</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mentation Strategies*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>New York, 2nd edition</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Zurich, ETH Zurich, 2007.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Athena Scientific, Belmont, Massachusetts, 3rd edition, 2005.</td>
<td>Massachusetts, 3rd edition,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Athena Scientific, Belmont, Massachusetts, 3rd edition, 2007.</td>
<td>Massachusetts, 3rd edition,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>tegies for minimising the risk of cascading failures in electricity</td>
<td><em>Critical Infrastructures,</em></td>
<td></td>
</tr>
</tbody>
</table>


Curriculum Vitae

Personal Data
Name: Michael Simon Guarisco
Date of birth: August 31st, 1977
Nationality: Swiss, citizen of Thalwil (ZH)
Marital status: Single

Education
1992 – 1997 High school (Swiss Matura, type C), Kantons-schule Heerbrugg (SG), Switzerland.
1998 – 2004 Studies in Mathematics at the Department of Mathematics, ETH Zurich, Switzerland. Graduated as Dipl. Math. ETH.
2005 – 2011 Doctoral student at the Institute for Operations Research, Department of Mathematics, ETH Zurich.

Experience
2005 – 2011 Member of the IT support group at the Institute for Operations Research, ETH Zurich.